

RESULT 33  
 Q9LNS7 PRELIMINARY; PRT; 651 AA.  
 AC Q9LNS7;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE T20H2.23 protein.  
 GN T20H2.23.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Sakano H., Vaysberg M., Lee J.M., Lenz C., Liu S., Pham P.,  
 RA Toriumi M., Yu G., Chin C., Chioi J., Choi E., Chung M., Gonzalez A.,  
 RA Howng B., Liu A., Altafi H., Brooks S., Buehler E., Chao Q., Conn L.,  
 RA Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B.,  
 RA Miranda M., Nguyen M., Palm C.J., Shin P., Southwick A., Davis R.W.,  
 RA Ecker J.R., Federspiel N.A., Theologis A.;  
 RT "The sequence of BAC T20H2 from Arabidopsis thaliana chromosome 1.";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC022472; AAF79914.1; -  
 SO SEQUENCE 651 AA; 73831 MW; 8AB2D146004F7896 CRC64;

Query Match 7.2%; Score 7; DB 10; Length 651;  
 Best Local Similarity 100.0%; Pred. No. 46;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 KKKRERK 34  
 Db 507 KKKRERK 513

RESULT 34  
 Q9V6S8 PRELIMINARY; PRT; 680 AA.  
 ID Q9V6S8;  
 AC Q9V6S8;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE CG13337 protein.  
 GN CG13337.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Beasley E.M.,  
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferliera S., Fleischmann W.,  
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Puri V., Reese M.G.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AF003817; AAF58344.1; -  
 DR FLYBASE; FBgn0033863; CG13337.  
 SO SEQUENCE 680 AA; 79758 MW; E10B28EFBE127FAD CRC64;

Query Match 7.2%; Score 7; DB 5; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 48;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 26 ERKKRRE 32  
 Db 461 ERKKRRE 467

RESULT 35  
 O15816 PRELIMINARY; PRT; 699 AA.  
 ID O15816;  
 AC O15816;  
 DT 01-JAN-1998 (TREMBlrel. 05, Created)  
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Myb2 protein.  
 GN MYB2.  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.  
 OX NCBI\_TaxID=44689;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AX3;  
 RX MEDLINE=98283920; PubMed=9620859;  
 RA Otsuka H., Van Haastert P.J.M.;  
 RT "A novel Myb homologue initiates Dictyostelium development by  
 induction of adenyl cyclase expression.";  
 RT Genes Dev. 12:1738-1748(1998).  
 RL -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 3 MYB-LIKE DOMAINS.  
 CC EMBL: AJ002383; CAA05357.1; -  
 DR HSSP; P06876; 1MBG.  
 DR InterPro; IPR001005; Myb DNA binding.  
 DR Pfam; PF00249; myb\_DNA\_binding; 3.  
 DR SMART; SMO0395; SANT; 3.  
 DR PROSITE; PS50090; MYB\_3; 3.  
 KW DNA-binding; Nuclear protein.  
 SO SEQUENCE 699 AA; 80381 MW; EA989DC77953F265 CRC64;

Query Match 7.2%; Score 7; DB 5; Length 699;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 KKKRERK 34  
 Db 409 KKKRERK 415

## RESULT 36

Q9RZF3 PRELIMINARY; PRT; 703 AA.  
AC Q9RZF3;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 20, Last annotation update)  
DE Nodulation protein-related protein.  
GN DRC0037.  
OS Deinococcus radiodurans.  
OC Plasmid Cpl.  
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;  
OC Deinococcaceae; Deinococcus.  
OX NCBI\_TaxID=1299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RI;  
RX MEDLINE=20036896; PubMed=10567266;  
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
RA Mofat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,  
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
RT Fraser C.M.;  
RT "Genome sequence of the radioresistant bacterium Deinococcus  
radiodurans RI."  
RL Science 286:1571-1577(1999).  
DR EMBL; AE01827; AAF12674.1; -.  
DR TIGR; DRC0037; -.  
DR InterPro; IPR003696; Carbohydr.  
DR Pfam; PF02543; CmcH\_Nodu; 1.  
KW Plasmid; Complete proteome.  
SQ SEQUENCE 703 AA; 77447 MW; 6ED24979D0ECD646 CRC64;

Query Match  
Best Local Similarity 7.2%; Score 7; DB 16; Length 703;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 GDSETL 18  
Db 165 GDSETL 171

## RESULT 37

Q9H1V1 PRELIMINARY; PRT; 772 AA.  
AC Q9H1V1;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE DJ127C7.1 (KIAA0250 protein).  
GN DJ127C7.1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lawlor S.;  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL137800; CAC19685.1; -.  
SQ SEQUENCE 772 AA; 85964 MW; FED3650A34447037 CRC64;

Query Match  
Best Local Similarity 7.2%; Score 7; DB 4; Length 772;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 SQTELRK 24  
Db 282 SQTELRK 288

## RESULT 38

O81868 PRELIMINARY; PRT; 777 AA.  
AC O81868;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE Hypothetical 87.4 kDa protein.  
GN T16H5.230 OR AT4G19870.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA De haan M., Maarse A.C., Grivell L.A., Bancroft I., Mewes H.W.,  
RA Mayer K., Schueller C., Bevan M.;  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA De Haan M., Maarse A.C., Grivell L.A., Mewes H.W., Lemcke K.,  
RA Mayer K.F.X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL024486; CA19704.1; -.  
DR EMBL; AL161551; CAB78989.1; -.  
DR InterPro; IPR001810; F-box.  
DR InterPro; IPR001798; Kelch.  
DR Pfam; PF00646; F-box; 2.  
DR Pfam; PF01344; Kelch; 4.  
DR PRINTS; PR00501; KELCHREPEAT.  
DR SMART; SM00256; FBOX; 2.  
DR PROSITE; PS50181; FBOX; 2.  
KW Hypothetical protein.  
SQ SEQUENCE 777 AA; 87450 MW; 1ABCC50FD6DED789 CRC64;

Query Match  
Best Local Similarity 7.2%; Score 7; DB 10; Length 777;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 ATRSHLG 92  
Db 455 ATRSHLG 461

## RESULT 39

Q96N31 PRELIMINARY; PRT; 799 AA.  
AC Q96N31;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE CDNA FLJ31473 fis, clone NT2NE2001530 (Fragment).  
GN Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,  
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,  
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,  
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,  
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,  
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,  
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,  
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.,  
RT "NEDO human cDNA sequencing project."  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK056035; BAB71079.1; -  
DR InterPro: IPR001440; TPR.  
DR SMART; SM00028; TPR; 1.  
FT NON\_TER 799 799  
SQ SEQUENCE 799 AA; 89725 MW; 5F3B5825B9BB01B0 CRC64;

Query Match 7.2%; Score 7; DB 4; Length 799;  
Best Local Similarity 100.0%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 SQTELRK 24  
Db 571 SQTELRK 577

## RESULT 40

O9BKN8 PRELIMINARY; PRT; 881 AA.  
ID O9BKN8  
AC O9BKN8;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Protein serine/threonine kinase-1.  
GN PSK1.  
OS Plasmodium falciparum.  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OX NCBI\_TaxID=5833;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-3D7A;  
RX MEDLINE=21203683; PubMed=11306117;  
RA Li J., Targett G.A., Baker D.A.;  
RT "Primary structure and sexual stage-specific expression of a LAMMER  
protein kinase of Plasmodium falciparum.";  
RL Int. J. Parasitol. 31:387-392(2001).  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
DR EMBL; AF104915; AAK38173.1;  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR002290; Ser\_thr\_pkinase.  
DR InterPro: IPR001245; Tyr\_pkinase.  
DR Pfam; PF00069; pkinase; 1.  
DR ProDom; PD000001; Euk\_pkinase; 1.  
DR SMART; SM00220; S\_TKC; 1.  
DR SMART; SM00219; TyrKC; 1.  
DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.  
KW ATP-binding; kinase; Serine/threonine-protein kinase; Transferase.  
SQ SEQUENCE 881 AA; 106705 MW; A4FB18FBB450B58 CRC64;

Query Match 7.2%; Score 7; DB 5; Length 881;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 KKERKKK 30  
Db 518 KKERKKK 524

## RESULT 41

O8RYX2 PRELIMINARY; PRT; 1054 AA.  
ID O8RYX2  
AC O8RYX2;  
DT 01-JUN-2002 (TREMBlrel. 21, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE OSJNB0066C06.5 protein.  
GN OSJNB0066C06.5.  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OX NCBI\_TaxID=39947;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN-CV. NIPPONBARE;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC  
clone:OSJNB0066C06.";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP003933; BAB90633.1;  
SQ SEQUENCE 1054 AA; 113161 MW; 912329E9D71C7D8E CRC64;

Query Match 7.2%; Score 7; DB 10; Length 1054;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 26 ERKKKRE 32  
Db 388 ERKKKRE 394

## RESULT 42

O95RH4 PRELIMINARY; PRT; 1283 AA.  
ID O95RH4  
AC O95RH4;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE LD29423p.  
GN CG4857 OR EG:EG0007.4 OR EG:EG0007.12.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BERKELEY;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,  
RA Yu C., Lewis S.E., Rubin G.M., Celnik S.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY061372; AAL28920.1;  
DR FLYBase; FBgn026083; CG4857.  
SQ SEQUENCE 1283 AA; 133625 MW; FFE0AFB9B7C17076 CRC64;

Query Match 7.2%; Score 7; DB 5; Length 1283;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 RKERKK 29  
Db 190 RKERKK 196

## RESULT 43

O21025 PRELIMINARY; PRT; 1286 AA.  
ID O21025  
AC O21025;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE F59A6.5 protein.  
GN F59A6.5.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for

```

RT Investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Nhan M.;
RT "The sequence of C. elegans cosmid F59A6."
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U41994; AAK31526.1; -.
DR InterPro; IPR001180; Cltron.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR001849; PH.
DR Pfam; PF00780; CNH; 1.
DR Pfam; PF00169; PH; 1.
DR SMART; SM00109; C1; 1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00233; PH; 1.
DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
SQ SEQUENCE 1286 AA; 148666 MW; 1402C3A80DC12BE4 CRC64;

Query Match
Best Local Similarity 7.2%; Score 7; DB 5; Length 1286;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 ASLGDSE 15
DB 317 ASLGDSE 323

RESULT 44
ID Q22944 PRELIMINARY; PRT; 1336 AA.
AC Q22944;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE C52B9.8 protein.
GN C52B9.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Nelson J.;
RT "The sequence of C. elegans cosmid C52B9."
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U64598; AAK39219.1; -.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00271; helicase_C; 1.
DR Pfam; PF00176; SNF2_N; 1.

```

```

DR PRINTS; PR01632; PQVCCALPHA1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELICC; 1.
KW ATP-binding; Helicase.
SQ SEQUENCE 1336 AA; 156127 MW; DBADE8CFAE0593B8 CRC64;

Query Match
Best Local Similarity 7.2%; Score 7; DB 5; Length 1336;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 QTELKK 25
DB 134 QTELKK 140

RESULT 45
ID Q03291 PRELIMINARY; PRT; 1435 AA.
AC Q03291;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Regulatory protein.
GN SPP41 OR YDR464W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E., Berno A.,
RA Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
RA Hunkeler-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
RA Mosedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
RA Winant A., Yelton M., Botstein D., Davis R.W.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Dietrich F.S.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Jia Y., Cherry J.M.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U33050; AAB64901.1; -.
DR SGD; S0002872; SPP41.
DR InterPro; IPR003903; UIM.
SQ SEQUENCE 1435 AA; 161596 MW; 7C18CF349FD647E8 CRC64;

Query Match
Best Local Similarity 7.2%; Score 7; DB 3; Length 1435;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 KKERKK 30
DB 695 KKERKK 701

```

Search completed: May 11, 2003, 20:12:09  
Job time : 56.5044 secs



GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 11, 2003, 16:05:46 ; Search time 2182.93 Seconds  
(without alignments)  
1293.204 Million cell updates/sec

Title: US-09-854-133-586  
Perfect score: 97  
Sequence: 1 EVEVSRDHASLDSETLSTQ.....LTGGCLPWATRSHLGRKCS 97

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4106486

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 60 summaries

Command line parameters:  
-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09854133/runat\_05052003\_174132\_704/app\_query.fasta.1.462  
-DB=GenEmbl -QEMT=fastap -SUFFIX=oligo.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=60  
-DOCALLIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=45 -MODE=LOCAL -OUTFMT=ptc  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09854133 @CGN\_1\_1\_1758 @runat\_05052003\_174132\_704 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database: GenEmbl:  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	97	100.0	337	6	AX321911	AX321911 Sequence
2	96	99.0	5981	6	AX321910	AX321910 Sequence
3	93	95.9	2239	6	AX321909	AX321909 Sequence
4	60	61.9	161280	9	AC093903	AC093903 Homo sapi
5	37	38.1	575	6	AX341060	AX341060 Sequence
6	37	38.1	596	6	AX351341	AX351341 Sequence
7	37	38.1	1683	9	AB042201	AB042201 Homo sapi
8	37	38.1	1861	9	AB026891	AB026891 Homo sapi
9	37	38.1	1874	9	AF200708	AF200708 Homo sapi
10	37	38.1	2000	9	AB040875	AB040875 Homo sapi
11	37	38.1	2155	9	BC012087	BC012087 Homo sapi
12	37	38.1	2482	9	AF252872	AF252872 Homo sapi
13	37	38.1	3144	5	HS277882	HS277882 Homo sapi
14	12	12.4	295	5	TCR292491	AJ292491 Triticum
15	12	12.4	681	4	CFA279982	AJ279982 Canis fam
16	12	12.4	41369	9	CH19F15314	AD000091 Homo sapi
17	12	12.4	68527	2	AC101290	AC101290 Mus muscu
18	12	12.4	73557	2	AC022364	AC022364 Homo sapi
19	12	12.4	76163	9	AP000960	AP000960 Homo sapi
20	12	12.4	100167	9	HS189613	AL121916 Human DNA
21	12	12.4	125267	4	AC090033	AC090033 Felis cat
22	12	12.4	125661	4	AC087731	AC087731 Felis cat
23	12	12.4	127475	2	AC105404	AC105404 Mus muscu
24	12	12.4	143786	2	AC099249	AC099249 Rattus no
25	12	12.4	145722	2	AC015503	AC015503 Homo sapi
26	12	12.4	151183	2	AC004932	AC004932 Homo sapi
27	12	12.4	152959	2	AC011275	AC011275 Homo sapi
28	12	12.4	153064	9	AC011492	AC011492 Homo sapi
29	12	12.4	156795	2	AC022290	AC022290 Homo sapi
30	12	12.4	159875	9	AC024023	AC024023 Homo sapi
31	12	12.4	163980	2	AC120209	AC120209 Felis cat
32	12	12.4	170513	2	AC023603	AC023603 Homo sapi
33	12	12.4	171506	2	AC120322	AC120322 Rattus no
34	12	12.4	172571	2	AC027484	AC027484 Homo sapi
35	12	12.4	173700	2	AC116823	AC116823 Mus muscu
36	12	12.4	173735	2	AC027063	AC027063 Homo sapi
37	12	12.4	175398	2	AC105990	AC105990 Mus muscu
38	12	12.4	175690	9	AC009989	AC009989 Homo sapi
39	12	12.4	177444	9	AC011468	AC011468 Homo sapi
40	12	12.4	178361	9	AC006042	AC006042 Homo sapi
41	12	12.4	178733	2	AC107305	AC107305 Homo sapi
42	12	12.4	179343	2	AC121904	AC121904 Mus muscu
43	12	12.4	180336	2	AC108976	AC108976 Rattus no
44	12	12.4	181445	9	AC026367	AC026367 Homo sapi
45	12	12.4	183011	2	AC118757	AC118757 Homo sapi
46	12	12.4	186298	9	AC090671	AC090671 Homo sapi
47	12	12.4	186298	2	AC093768	AC093768 Homo sapi
48	12	12.4	188690	2	AC121187	AC121187 Rattus no
49	12	12.4	189169	2	AC122690	AC122690 Homo sapi
50	12	12.4	193171	2	AC121771	AC121771 Mus muscu

51	12	12.4	193887	9	AF313041	AF313041	Homo sapi
C 52	12	12.4	193957	2	AC090198	AC090198	Homo sapi
53	12	12.4	194407	2	AC068845	AC068845	Homo sapi
54	12	12.4	194851	9	AC079264	AC079264	Homo sapi
55	12	12.4	195782	9	AC074389	AC074389	Homo sapi
56	12	12.4	196710	2	AC126470	AC126470	Homo sapi
C 57	12	12.4	197078	2	AL845536	AL845536	Mus muscu
C 58	12	12.4	203734	2	AC118229	AC118229	Mus muscu
C 59	12	12.4	205914	2	AC131101	AC131101	Mus muscu
C 60	12	12.4	206880	2	AC108392	AC108392	Mus muscu

## ALIGNMENTS

RESULT	1
AX321911	
LOCUS	AX321911
DEFINITION	Sequence 442 bp from Patent WO0172295.
ACCESSION	AX321911
VERSION	AX321911.1
KEYWORDS	GI:17906521
SOURCE	.
ORGANISM	human.
	Homo sapiens
PAT	15-DEC-2001
linear	DNA

REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES
1	Reed, S.G., Lodes, M.J., Mohamath, R., Secrist, H., Benson, D.R., Indrias, C.Y., Henderson, R.A., Fling, S.P., Algate, P.A., Elliot, M., Mannion, J. and Kalos, M.D.	Compositions and methods for the therapy and diagnosis of lung cancer	Patent: WO 0172295-A 442 04-OCT-2001;	CORIXA CORPORATION (US)
				Location/Owner/Accession Number

FEATURES	source
Location/Qualifiers	1. .337
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
103 a	60 c 93 g . 81 t
BASE COUNT	
ORIGIN	

Alignment Scores:	
Pred. No.:	4.23e-94
Score:	97.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	6
Length:	337
Matches:	97
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

03 03-834-133-586 (1-97) x AX321911 (1-337)

QY	1	20
Db	5	64
QY	21	64
Db	65	124
QY	41	60
Db	125	184
QY	61	80
Db	185	244
QY	81	97
Db	245	295
RESULT 2		
AX321910	5981 bp	DNA linear PAT 15-DEC-2001

DEFINITION	Sequence 441 from Patent WO0172295
ACCESSION	AX321910
VERSION	AX321910.1
KEYWORDS	GI:17906518
SOURCE	human.

REFERENCE  
AUTHORS  
1 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
Reed, S.G., Lodes, M.J., Mohamath, & Scardina, R. 2003. The evolution of the primate visual system: a review of the fossil record and the molecular data.

TITLE	Author
Compositions and methods for the therapy and diagnosis of lung cancer	Indirias, C.Y., Henderson, R.A., Fling, S.P., Algate, P.A., Elliot, M., Mannion, J. and Kalos, M.D.

JOURNAL Patent: WO 0172295-A 441 04-OCT-2001;  
CORIXA CORPORATION (US)

BASE COUNT	1858	a	1029	c	1098	g	1996	t
ORIGIN	/organism="Homo sapiens" /db_xref="taxon:9606"							

Alignment Scores:	
Pred. No.:	5.97e-92
Score:	96.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	98.97%
DB:	6
Length:	5981
Matches:	96
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

05-09-854-133-586 (1-97) x AX321910 (1-5981)

QY	2	ValGluValSerArgAspHisAlaSerLeuGlyAspSerCyluThrLeuSerGlnThrGlu	21
Db	3	GTTGAAGTGAAGCAGAGATCATGCCAGCCTGGGTGACAGTGAGACTGTCTCAACAGAA	
QY	22	LeuArgLysLysGluArgLysLysLysArgGluArgLysPheGlnAlaAsnCysGlyIle	41
Db	63	TTAAGCAAAAAAGAAAGAAAAAGAGAGAGAGAAATCCAGGCCAATGTGGCATA	
QY	42	AspPheIleIlePheTrpIlePheTrpIleLeuPheSerHisIleTrpIleGlnGlu	61
Db	123	GATTTATCATATTTCTGGATTTTGGATTCTTTGTTCTTCATCATCGATTCCAGAA	182
QY	62	SerLeuLeuCysProProSerProLysGluValThrCysArgGluMetLeuThrGlyGly	81
Db	183	AGCCGTGTGTGCCACCATCTCCAAAGAGGTTACCTGCAGGGAATTTAAACGGGAGGC	242
QY	82	CysLeuProTrpAlaThrArgSerHisLeuGlyArgArgLysCysSer	97
Db	243	TGCCTTCCCTGGGCAACAAGAGCCACCTGGGACAGAGAAAGTGCCAGC	290

RESULT 3			
AX321909	AX321909	2239 bp	DNA
LOCUS			
DEFINITION	Sequence	440 from Patent WO0172295.	linear
ACCESSION	AX321909		
VERSION	AX321909.1	GI:17906515	
KEYWORDS			
			PAT 15-DEC-2001

ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1
AUTHORS	Reed,S.G., Lodes,M.J., Mohamath,R., Secrist,H., Benson,D.R., Indirias,C.Y., Henderson,R.A., Fling,S.P., Algate,P.A., Elliot,M. Mannion,J. and Kalos,M.D.
TITLE	Compositions and methods for the therapy and diagnosis of lung cancer
JOURNAL	Patent: WO 01/22295-A 440 04-OCT-2001; CORIXA CORPORATION (US)
FEATURES	Location/Qualifiers

source 1..2239  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 619 a 444 c 493 g 683 t

ORIGIN

Alignment Scores: 3.94e-89 Length: 2239  
Pred. No.: 93.00 Matches: 93  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 95.88% Indels: 0  
Query Match: 6 Gaps: 0

US-09-854-133-586 (1-97) x AX321909 (1-2239)

OY 1 GUAUAGUValserARgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThr 20  
|||||  
Db 2 GAGGTGAAGTGAAGCAGAGATCAATGCCAGCTGGGTGACAGTGAAGACTCTGTCTCAACA 61  
OY 21 GUAUArgLysLysGUAArgLysLysArgGUAArgLysPheGlnAlaAsnCysGly 40  
|||||  
Db 62 GAATTAAGGAAAAAGAAAGAAAGAAAGAGAGAGAGAAATTCACAGGCCAATTGTGGC 121  
OY 41 ILeaspHeIleIlePheTrpIlePheTrpIleLeuLeuPheSerHisIleTrpIleGln 60  
|||||  
Db 122 ATAGATTTCATATCTGATTTTTCGATTTTTCATCTTCATCTGATTCAG 181  
OY 61 GUAUArgLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 80  
|||||  
Db 182 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGAGGTTACCTGACAGGAAATGTTAAGGGA 241  
OY 81 GUAUArgLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 93  
|||||  
Db 242 GGCTGCTTCCCTGGGCAACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 280

RESULT 4  
AC093903/c 161280 bp DNA linear PRI 01-MAR-2002  
LOCUS Homo sapiens BAC clone RP11-733C7 from 4, complete sequence.  
DEFINITION AC093903 AC055827  
AC093903.3 GI:15920156  
VERSION  
KEYWORDS HTG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 161280)  
TITLE Sultston, J.E. and Waterston, R.  
JOURNAL Toward a complete human genome sequence  
MEDLINE Genome Res. 8 (11), 1097-1108 (1998)  
PUBMED 99063792  
9847074  
REFERENCE  
AUTHORS 2 (bases 1 to 161280)  
TITLE Radionenko, M. and Kozlowski, A.  
JOURNAL The sequence of Homo sapiens BAC clone RP11-733C7  
AUTHORS Unpublished (2001)  
TITLE 3 (bases 1 to 161280)  
JOURNAL Waterston, R.H.  
AUTHORS Direct Submission  
TITLE Submitted (10-SEP-2001) Genome Sequencing Center, Washington  
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
REFERENCE  
AUTHORS 4 (bases 1 to 161280)  
TITLE Waterston, R.H.  
JOURNAL Direct Submission  
AUTHORS Submitted (04-OCT-2001) Genome Sequencing Center, Washington  
JOURNAL University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
REFERENCE  
AUTHORS 5 (bases 1 to 161280)  
TITLE Waterston, R.  
JOURNAL Direct Submission  
AUTHORS Submitted (01-MAR-2002) Department of Genetics, Washington

## COMMENT

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Oct 4, 2001 this sequence version replaced gi:15625016.

## ----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc>

Contact: [sapiens@wustl.edu](mailto:sapiens@wustl.edu)

----- Summary Statistics

Center project name: H\_NH0733C07

Drafting Center: WIBR

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

## MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

## SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACE3.6

## NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is RP11-310A13. Actual start of this clone is at base position 1 of RP11-733C7; actual end is at base position 161280 of RP11-733C7.

Data from AC09792 was used to finish the clone, AC055827.

The sequence of AC055827 has been incorporated into AC093903.

## FEATURES

## source

Location/Qualifiers

1..161280

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="4"

/map="4"

/clone="RP11-733C7"

/clone\_lib="RPCI-11"

148..304

/rpt\_family="L1"

306..406

/rpt\_family="(GGAA)n"

471..525

/rpt\_family="AT-rich"

530..782

/rpt\_family="L1"

685..712

/rpt\_family="(TA)n"

1383..1472

/rpt\_family="GA-rich"

1697..1836

/rpt\_family="L1"

1837..1956

/rpt\_family="L1"

repeat\_region

repeat\_region

repeat\_region

```

repeat_region 3005. .3221
/rpt_family="MIR"
repeat_region 4013. .4321
/rpt_family="Alu"
repeat_region 4295. .4325
/rpt_family="(A)n"
repeat_region 5625. .5960
/rpt_family="L1"
repeat_region 5938. .5982
/rpt_family="A-rich"
repeat_region 6062. .6195
/rpt_family="MIR"
repeat_region 6388. .6462
/rpt_family="L2"
repeat_region 6546. .6868
/rpt_family="L1"
misc_feature 6866. .7318
/note="similar to EST C06175 (NID:g1502951)"
repeat_region 6866. .6898
/rpt_family="AT-rich"
misc_feature 7134. .7385
/note="similar to EST BG989897 (NID:g14393967)"
misc_feature 7777. .7918
/note="similar to EST BG388107 (NID:g13281553)"
repeat_region 8483. .8522
/rpt_family="AT-rich"
repeat_region 8721. .9018
/rpt_family="Alu"
repeat_region 9459. .9741
/rpt_family="L2"
repeat_region 10217. .10246
/rpt_family="AT-rich"
repeat_region 10304. .10385
/rpt_family="MIR"
repeat_region 10835. .10878
/rpt_family="AT-rich"
misc_feature 11861. .11987
/note="similar to EST BG388107 (NID:g13281553)"
misc_feature 11873. .11987
/note="similar to EST BF286052 (NID:g11217122)"
repeat_region 12516. .12539
/rpt_family="(T)n"
repeat_region 12929. .13232
/rpt_family="Alu"
repeat_region 13237. .13274
/rpt_family="(TA)n"
repeat_region 13275. .13304
/rpt_family="(T)n"
repeat_region 13696. .13838
/rpt_family="MIR"
repeat_region 14127. .14153
/rpt_family="AT-rich"
repeat_region 14132. .14436
/rpt_family="Alu"
repeat_region 15823. .16117
/rpt_family="Alu"
misc_feature 17329. .17894
/note="similar to EST BG388107 (NID:g13281553)"
misc_feature 17329. .17618
/note="similar to EST BF286052 (NID:g11217122)"
misc_feature 17489. .17611
/note="similar to EST AI313891 (NID:g4029010)"
misc_feature 17548. .17869
/note="similar to EST BF095483 (NID:g10901193)"
misc_feature 17772. .18329
/note="similar to EST BG181259 (NID:g13703031)"
misc_feature 17845. .17869
/rpt_family="(CTG)n"
repeat_region 18792. .18828
/rpt_family="AT-rich"
repeat_region 20734. .20784
/rpt_family="MERL-type"
repeat_region 20832. .20944

```

```

repeat_region /rpt_family="MERL-type"
20948. .21964
repeat_region /rpt_family="L1"
21964. .21986
repeat_region /rpt_family="AT-rich"
21965. .22276
repeat_region /rpt_family="Alu"
22277. .22323
repeat_region /rpt_family="L1"
22341. .22681
repeat_region /rpt_family="L1"
22685. .22824
repeat_region /rpt_family="Alu"
22805. .22824
repeat_region /rpt_family="(A)n"
22826. .23575
repeat_region /rpt_family="L1"
23591. .23735
repeat_region /rpt_family="MERL-type"

```

## Alignment Scores:

```

Pred. No.: 2.03e-52 Length: 161280
Score: 60.00 Matches: 60
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 61.86% Indels: 0
DB: 9 Gaps: 0

```

US-09-854-133-586 (1-97) x AC093903 (1-161280)

```

QY 1 GLUVALGLVALSERARGSPHISALASERLEUGLYASPSERGLUTHREUSERGLNTHR
Db 27755 GAGGTGAGTGAAGCAGAGATCATGCCCTGGGTGACAGTGAAGTCTGTCTCAACA
QY 21 GLUVALGLVALSERARGSPHISALASERLEUGLYASPSERGLUTHREUSERGLNTHR
Db 27695 GAGGTGAGTGAAGCAGAGATCATGCCCTGGGTGACAGTGAAGTCTGTCTCAACA
QY 41 ILEASPHEILEILEPHEITRIPILEULEUPHESERHISHISTRIPILEGIN
Db 27635 ATAGATTTTATCATATCTGATTTTGGATTTCTTTCTCATCATCATCATCATCAT
RESULT 5
AX341060 575 bp DNA linear PAT 10-JAN-2002
LOCUS Sequence 1307 from Patent WO0196388.
DEFINITION AX341060
ACCESSION AX341060
VERSION AX341060.1 GI:18137042
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Jiang,Y., Harlocker,S.L. and Secrist,H.
TITLE Compositions and methods for the therapy and diagnosis of colon
cancer
JOURNAL Patent: WO 0196388-A 1307 20-DEC-2001;
CORIXA CORPORATION (US)
FEATURES
source Location/Qualifiers
1..575
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 151 a 122 c 153 g 141 t 8 others
ORIGIN

```

## Alignment Scores:

```

Pred. No.: 4.38e-30 Length: 575
Score: 37.00 Matches: 37
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 38.14% Indels: 0
DB: 6 Gaps: 0

```



```

source
1: 1000
/organism="Homo sapiens"
/db_xref="taxon:9606"

```

gene  
/cell\_type="fibroblast"  
1.1861  
/gene="xct"  
236.1741  
/gene="xct"

/product\_start=1  
/product="cystine/glutamate transporter"  
/protein\_id="BAA82628.1"  
/db\_xref="GI:5668545"

BASE COUNT 485 a 402 c 421 g 553 t  
ORIGIN  
/translation="MVRKPVVSTISKGGYLOGNVNGLPSLGNKEPPGQEKVQLKRV  
TLRGVSIITIGTIGAGIFISPKGVLONTGVSMSLTITWVCGVLSFGALSYAELGT  
TIKSGGHYTYILEVFGPLPAFVRVWVLLIRPATAVISLAFGRYLEPFIQCEI  
PELAIKLITAVGITVVMVNLNSVSWASARIOIFLFECKLTAILIIVPGVMOLIKGT  
ONFKDAFSGRDSITRLPLAFYGMAYAGWFLNFVEEVENPEKTIPLAICISMAI  
VTIGVLTNAVAYFTTINAEELLISNAVATFSERLGNFSLAVPIFVALSCFGSMNG  
VEAVSRLEFYVASREGHLPEILSMIHVRKHTPLPAVIVLHPLTMIMLFSGLDLSLNL  
SEARWLEFLAVAGLIYLRKCPDMHRPEKVPLEIPALFSFTCLFMVALSLYSDPEST  
GIGFVITLTGPAYYLFITMDKPRWRIMSEKITRTLQIILEVVPEDKL"

## Alignment Scores:

Pred. No.: 1.22e-29 Length: 1861  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.14% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x AB026891 (1-1861)

OY 61 GluSerLeuLeuCySPProSerProlysgluValThrCysArggluMetLeuThrGly 80  
|||||  
Db 243 GAAAGCCTGTGTGTCCACCATCTCCAAGAGAGGTTACCTGCAGGAAATGTTAACGGGA 302  
|||||

OY 81 GlyCysLeuProTrrPalatThrArgSerHisLeuGlyArgArgLysCysSer 97  
|||||  
Db 303 GGCTGCCTTCCCTGGGCAACAGAGCCACCTGGCGAGAGAAAGTGCAGC 353  
|||||

## RESULT 9

AF200708 1874 bp mRNA linear PRI 01-DEC-2000  
LOCUS Homo sapiens calcium channel blocker resistance protein CCBRI mRNA,  
DEFINITION complete cds.  
ACCESSION AF200708  
VERSION AF200708.1 GI:11493651

## KEYWORDS

AF200708.1 GI:11493651

## SOURCE

Homo sapiens.  
Homo sapiens.

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.

## REFERENCE

1 (bases 1 to 1874)  
Conklin,D.S. and Beach,D.H.  
CCBR1, novel CD98 light chain implicated in redox control and  
calcium signaling  
Unpublished

## AUTHORS

2 (bases 1 to 1874)  
Conklin,D.S. and Beach,D.H.  
Direct Submission

## JOURNAL

Submitted (01-NOV-1999) Cold Spring Harbor Labs, 1 Bungtown Rd,  
Cold Spring Harbor, NY 11724, USA

## FEATURES

Location/Qualifiers  
1.1874

## CDS

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_line="fibrosarcoma HT1080"  
246.1751  
/note="X-c transporter; CD98 light chain"  
/codon\_start=1  
/product="calcium channel blocker resistance protein  
CCBR1"  
/protein\_id="AAG35592.1"  
/db\_xref="GI:11493652"

/translation="MVRKPVVSTISKGGYLOGNVNGLPSLGNKEPPGQEKVQLKRV  
TLRGVSIITIGTIGAGIFISPKGVLONTGVSMSLTITWVCGVLSFGALSYAELGT  
TIKSGGHYTYILEVFGPLPAFVRVWVLLIRPATAVISLAFGRYLEPFIQCEI  
PELAIKLITAVGITVVMVNLNSVSWASARIOIFLFECKLTAILIIVPGVMOLIKGT  
ONFKDAFSGRDSITRLPLAFYGMAYAGWFLNFVEEVENPEKTIPLAICISMAI  
VTIGVLTNAVAYFTTINAEELLISNAVATFSERLGNFSLAVPIFVALSCFGSMNG  
VEAVSRLEFYVASREGHLPEILSMIHVRKHTPLPAVIVLHPLTMIMLFSGLDLSLNL  
SEARWLEFLAVAGLIYLRKCPDMHRPEKVPLEIPALFSFTCLFMVALSLYSDPEST  
GIGFVITLTGPAYYLFITMDKPRWRIMSEKITRTLQIILEVVPEDKL"

## Alignment Scores:

Pred. No.: 1.22e-29 Length: 1874  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.14% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x AF200708 (1-1874)

OY 61 GluSerLeuLeuCySPProSerProlysgluValThrCysArggluMetLeuThrGly 80  
|||||  
Db 253 GAAAGCCTGTGTGTCCACCATCTCCAAGAGAGGTTACCTGCAGGAAATGTTAACGGGA 312  
|||||

OY 81 GlyCysLeuProTrrPalatThrArgSerHisLeuGlyArgArgLysCysSer 97  
|||||  
Db 313 GGCTGCCTTCCCTGGGCAACAGAGCCACCTGGCGAGAGAAAGTGCAGC 363  
|||||

## RESULT 10

AB040875 2000 bp mRNA linear PRI 03-APR-2001  
LOCUS Homo sapiens hXCT mRNA for cystine/glutamate exchanger, complete  
DEFINITION cds.  
ACCESSION AB040875  
VERSION AB040875.1 GI:13516845

## KEYWORDS

AB040875.1 GI:13516845

## SOURCE

Homo sapiens adult CDNA to mRNA, clone\_lib:brain CDNA library  
clone:hXCT.  
Homo sapiens

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.

## REFERENCE

1 (sites)  
Kim,J.Y., Chairoungdua,A., Cha,S.H., Segawa,H., Matsuo,H.,  
Kim,D.K., Endou,H. and Kanai,Y.  
Human cystine/glutamate exchanger: cDNA cloning and upregulation by  
oxidative stress in glioma cells  
Unpublished

## AUTHORS

2 (bases 1 to 2000)  
Kanai,Y.  
Direct Submission

## JOURNAL

Submitted (01-APR-2000) Yoshikatsu Kanai, Kyorin University School  
of Medicine, Department of Pharmacology and Toxicology; 6-20-2  
Shinkawa, Mitaka, Tokyo 181-8611, Japan  
(E-mail:ykanai@kyorin-u.ac.jp, Tel:+81-422-47-5511(ex.3453),  
Fax:+81-422-79-1321)

## FEATURES

Location/Qualifiers  
1.2000

## CDS

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="hXCT"  
/clone\_lib="brain CDNA library"  
/dev\_stage="adult"  
1.2000  
/gene="hXCT"  
136.1620  
/note="Na+-independent amino acid transporter  
transmembrane protein"  
/codon\_start=1  
/product="cystine/glutamate exchanger"  
/protein\_id="BAB40574.1"

/db\_xref="GI:13516846"  
/translation="MVRKPVSTISKGYLQGNVNGRLPSLGNKEPPGQEKVQLKRV  
TLRGVSIITIGITGAGIFISPKVLQNTGVSMSLTITWVCGVLSFGALSYAELGT  
TIKSGGHYTYILEVFGPLPAFVRWVELLIRPATAVISLAFGRILEPFIQCEI  
PELAIKLITAVGITVVMVNSVSWASARIQIFLTFCKLITAILIIVPGVQLIKQGT  
QNFKAESGRDSSITRLPLAFYGYMAYAGVFLNFTVEEVENPEKTIPLAICISMAI  
VTIGVLTNVAVFTTINAEELLSNAVAVTFSERLGNFSLAVPIFVALSCGSMNG  
VEAVSRLEFYVASREGHLEPILSMIHVRKHTPLPAVIVLHPLTLMILFSGDLSLNL  
SEARWLEIGLAVAGLITLYRKCPDMHPRFKVPLFIPALFSTCLFMVALSLYSDPFST  
GIGFVITLTVGPAYLFIWDRKPRWFRIMSGFALMPAQACDM"

BASE COUNT 505 a 418 c 435 g 642 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.3e-29 Length: 2000  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.14% Indels: 0  
DB: Gaps: 9

US-09-854-133-586 (1-97) x AB040875 (1-2000)

QY 61 GluserleuLeuCySPProSerProLysGluValThrcysArgGluMetLeuThrgly 80  
DB 143 GAAAGCCTGTGTGTCTCCACCATCTCCAAAGAGGTTACCTGCAGGGAATGTAAACGGGA 202

QY 81 GlyCysLeuProTrpAlaThrArgSerHisLeuGlyArgArgLysCysSer 97  
DB 203 GGCTGCCCTTCCCTGGCAACAGAGCCACTGCGCAGAGAAAGTGCAGC 253

RESULT 11  
BC012087 2155 bp mRNA linear PRI 06-AUG-2001  
LOCUS Homo sapiens, similar to solute carrier family 7, (cationic amino  
acid transporter, y+ system) member 11, clone MGC:20026  
IMAGE:4562994, mRNA, complete cds.

ACCESSION BC012087  
VERSION BC012087.1 GI:15082351  
KEYWORDS MGC.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 2155)  
Strausberg, R.  
Direct Submission  
Submitted (02-AUG-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: DCTD/DTP  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Genome Sequence Centre,  
BC Cancer Agency, Vancouver, BC, Canada  
[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)  
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,  
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,  
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo  
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven  
Ness, Pawan Pandoh, Anna-Lisa Prabh, Parvaneh Saeedi, Jacqueline  
Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,  
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,  
George Yang, Scott Zuyderduyn, Marco Marra.

REMARK  
COMMENT  
Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 29 Row: 9 Column: 1  
This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 5668544.  
Location/Qualifiers  
1. 2155  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MGC:20026 IMAGE:4562994"  
/tissue\_type="Kidney, renal cell adenocarcinoma"  
/clone\_id="NIH\_MGC\_14"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
161. 1666  
/codon\_start=1  
/product="Similar to solute carrier family 7, (cationic  
amino acid transporter, y+ system) member 11"  
/protein\_id="AAH12087.1"  
/db\_xref="GI:15082352"  
/translation="MVRKPVSTISKGYLQGNVNGRLPSLGNKEPPGQEKVQLKRV  
TLRGVSIITIGITGAGIFISPKVLQNTGVSMSLTITWVCGVLSFGALSYAELGT  
TIKSGGHYTYILEVFGPLPAFVRWVELLIRPATAVISLAFGRILEPFIQCEI  
PELAIKLITAVGITVVMVNSVSWASARIQIFLTFCKLITAILIIVPGVQLIKQGT  
QNFKAESGRDSSITRLPLAFYGYMAYAGVFLNFTVEEVENPEKTIPLAICISMAI  
VTIGVLTNVAVFTTINAEELLSNAVAVTFSERLGNFSLAVPIFVALSCGSMNG  
VEAVSRLEFYVASREGHLEPILSMIHVRKHTPLPAVIVLHPLTLMILFSGDLSLNL  
SEARWLEIGLAVAGLITLYRKCPDMHPRFKVPLFIPALFSTCLFMVALSLYSDPFST  
GIGFVITLTVGPAYLFIWDRKPRWFRIMSEKTRTLQILLEVPEEDKL"

BASE COUNT 599 a 443 c 452 g 661 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.38e-29 Length: 2155  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.14% Indels: 0  
DB: Gaps: 9

US-09-854-133-586 (1-97) x BC012087 (1-2155)

QY 61 GluserleuLeuCySPProSerProLysGluValThrcysArgGluMetLeuThrgly 80  
DB 168 GAAAGCCTGTGTGTCTCCACCATCTCCAAAGAGGTTACCTGCAGGGAATGTAAACGGGA 227

QY 81 GlyCysLeuProTrpAlaThrArgSerHisLeuGlyArgArgLysCysSer 97  
DB 228 GGCTGCCCTTCCCTGGCAACAGAGCCACTGCGCAGAGAAAGTGCAGC 278

RESULT 12  
AF252872 2482 bp mRNA linear PRI 02-MAY-2001  
LOCUS Homo sapiens cystine/glutamate transporter xCT mRNA, complete cds.  
ACCESSION AF252872  
VERSION AF252872.1 GI:13924719  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 2482)  
Chaney, C.D., Kekuda, R., Wang, H., Huang, W., Prasad, P.D., Smith, S.B.  
and Ganapathy, V.  
Structure, Function and Regulation of Human Cystine/Glutamate  
Transporter in Retinal Pigment Epithelial Cells  
Unpublished  
2 (bases 1 to 2482)  
Wang, H., Prasad, P.D. and Ganapathy, V.  
Direct Submission  
Submitted (05-APR-2000) Biochemistry & Molecular Biology, Medical  
College of Georgia, 1120 15th Street, Augusta, GA 30912, USA  
Location/Qualifiers  
1. 2482  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tissue\_type="placenta"

CDS

232..1737  
/codon\_start=1  
/product="cystine/glutamate transporter xct"  
/protein\_id="AAK4911.1"  
/db\_xref="GI:13924720"  
/translation="MVRPVYSTISKGYLQGNVNGRLPSLNKEPPOGKVKQLKRY  
TIKSGGHVYILEFVGLPAFVRVWVLLIRPAATAVISLAFGRVILEPFIQCEI  
PELAIKLITAVGITVVMVNSVMSVARSARIQIFLTFCKLTAIIIVGMOLIKGOT  
QNEKDAFSGRDSITRLPLAFYGYMAYAGWFLNFEVEENPEKTIPLAICISMAI  
VEAVSRLEFYVASREGLPELISMIVHVKHTPLPAVIVLHPLTMIMLFSGLDLSLNL  
SFARWLEFLAVAGLILYLRKCPDMHRPEKVPFLIPALFSTCLFMVALSLSDPFS  
GIGFVITLGVPAVYFLIWDKPRWFRIMSEKITRILQIILEVPEEDKL"

BASE COUNT 735 a 484 c 521 g 742 t

ORIGIN

Alignment Scores:  
Pred. No.: 1.56e-29  
Score: 37.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 38.14%  
DB: 9

Length: 2482  
Matches: 37  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x AF252872 (1-2482)

QY 61 GluSerLeuLeuCySPProSerProLySGluValThrCysArgGluMetLeuThrGly 80  
|||||  
Db 239 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGAGGTACCTGCAGGGAATGTTAACGGGA 298  
|||||

QY 81 GlyCysLeuProTrpAlaThrArgSerHisLeuGlyArgArgLysCysSer 97  
|||||  
Db 299 GGCTGCCCTTCCCTGGGCAACAAGAGCCACCTGGCAGAGAGAAAGTGACG 349  
|||||

RESULT 13  
HSA277882  
LOCUS HSA277882 3144 bp mRNA linear PRI 04-JAN-2002  
DEFINITION Homo sapiens mRNA for cystine/glutamate transporter (XCT gene).  
ACCESSION AJ277882  
VERSION AJ277882.1 GI:18073361  
KEYWORDS cystine/glutamate transporter; XCT gene.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
AUTHORS Borsani, G., Manzoni, M., Palacin, M., Pineda, M. and Gasol, E.  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 3144)  
AUTHORS Bassi, M.T.  
TITLE Direct Submission  
JOURNAL Submitted (11-MAY-2000) Bassi M.T., Telethon Institute of Genetics  
and Medicine, Via Olgettina 58, 20132 Milan, ITALY

FEATURES  
source  
1..3144  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="4"  
/map="4q28-q32"  
/cell\_line="NT2, undifferentiated teratocarcinoma cell  
1..3128  
/gene="XCT"  
1..1506  
/gene="XCT"  
/codon\_start=1  
/product="cystine/glutamate transporter"  
/protein\_id="CAC81905.1"  
/db\_xref="GI:18073362"  
/translation="MVRPVYSTISKGYLQGNVNGRLPSLNKEPPOGKVKQLKRY  
TIKSGGHVYILEFVGLPAFVRVWVLLIRPAATAVISLAFGRVILEPFIQCEI  
TIKSGGHVYILEFVGLPAFVRVWVLLIRPAATAVISLAFGRVILEPFIQCEI

CDS  
gene  
1..1506  
/gene="XCT"  
1..1506  
/gene="XCT"  
/codon\_start=1  
/product="cystine/glutamate transporter"  
/protein\_id="CAC81905.1"  
/db\_xref="GI:18073362"  
/translation="MVRPVYSTISKGYLQGNVNGRLPSLNKEPPOGKVKQLKRY  
TIKSGGHVYILEFVGLPAFVRVWVLLIRPAATAVISLAFGRVILEPFIQCEI

repeat\_region  
polyA\_site  
BASE COUNT 912 a 602 c 649 g 981 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.92e-29  
Score: 37.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 38.14%  
DB: 9

Length: 3144  
Matches: 37  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x HSA277882 (1-3144)

QY 61 GluSerLeuLeuCySPProSerProLySGluValThrCysArgGluMetLeuThrGly 80  
|||||  
Db 8 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGAGGTACCTGCAGGGAATGTTAACGGGA 67  
|||||

QY 81 GlyCysLeuProTrpAlaThrArgSerHisLeuGlyArgArgLysCysSer 97  
|||||  
Db 68 GGCTGCCCTTCCCTGGGCAACAAGAGCCACCTGGCAGAGAGAAAGTGACG 118  
|||||

RESULT 14  
TCR292491  
LOCUS TCR292491 295 bp DNA linear VRT 16-MAY-2002  
DEFINITION Triturus cristatus microsatellite DNA, locus Tcr36.  
ACCESSION AJ292491  
VERSION AJ292491.1 GI:20798953  
KEYWORDS microsatellite; repetitive element.  
SOURCE warty newt.  
ORGANISM Triturus cristatus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Amphibia; Batrachia; Caudata; Salamandroidea; Salamandridae;  
Triturus.

REFERENCE  
AUTHORS Krupa, A.P., Jehle, R., Dawson, D.A., Gentle, L.K., Gibbs, M.,  
Arntzen, J.W. and Burke, T.  
JOURNAL Microsatellite loci in the crested newt (Triturus cristatus) and  
their utility in other newt taxa  
REFERENCE 2 (bases 1 to 295)  
AUTHORS Krupa, A.P.  
TITLE Direct Submission  
JOURNAL Submitted (18-JUL-2000) Krupa A.P., Animal and Plant Sciences,  
University of Sheffield, Alfred Denny Building, Western Bank,  
Sheffield, South Yorkshire, S10 2TN, UNITED KINGDOM

FEATURES  
source  
1..295  
/organism="Triturus cristatus"  
/db\_xref="taxon:8323"  
/clone="TC7363"  
/note="locus Tcr36"

satellite  
BASE COUNT 156 a 26 c 69 g 43 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 0.000952  
Score: 12.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 12.37%  
DB: 5

Length: 295  
Matches: 12  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0



US-09-854-133-586 (1-97) x TCR292491 (1-295)

OY 23 ArgLysLysGluArgLysLysLysArgGluArgLys 34  
 |||||  
 Db 99 AGAAGAGAAAGAAAGAAAGAAAGAGAGAAG 134

RESULT 15  
 CFA279982 681 bp DNA linear MAM 18-AUG-2000

LOCUS CFA279982 Canis familiaris microsatellite DNA, clone DTRcn.31.

DEFINITION AJ279982

ACCESSION AJ279982.1 GI:9856989

VERSION microsatellite; repetitive DNA; satellite.

KEYWORDS dog.

SOURCE Canis familiaris

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

REFERENCE 1 (bases 1 to 681)  
 Turet, L., Kessler, J.L., Bentolila, S., Faure, S., Bach, J.M.,  
 Weissenbach, J. and Panthier, J.J.

AUTHORS Assignment of highly polymorphic markers on a canine purebred pedigree

TITLE Mamm. Genome 11 (8), 703-705 (2000)

JOURNAL Mamm. Genome 11 (8), 703-705 (2000)

MEDLINE 20380834

PUBMED 10920245

REFERENCE 2 (bases 1 to 681)  
 Weissenbach, J.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-JAN-2000) Weissenbach J., Genoscope, Centre National de Sequencage, 2 rue Gaston Cremieux, 91006 Evry cedex, FRANCE

FEATURES  
 source  
 1. .681  
 /organism="Canis familiaris"  
 /strain="Beagle"  
 /db\_xref="taxon:9615"  
 /clone="DTRcn.31"  
 /tissue\_type="blood"  
 145. .164  
 /note="STS primer DTRcn.31\_f (AJ287256)"  
 /PCR\_conditions="30 x (94OC, 30sec; 55OC, 40sec; 72OC, 40sec)"  
 165. .469  
 /note="polymorphic tetranucleotide repeat"  
 /rpt\_type=TANDEM  
 complement(483. .505)  
 /note="STS primer DTRcn.31\_r (AJ287257)"  
 /PCR\_conditions="30 x (94OC, 30sec; 55OC, 40sec; 72OC, 40sec)"

primer\_bind

satellite

primer\_bind

BASE COUNT 342 a 60 c 145 g 132 t 2 others  
 ORIGIN

Alignment Scores:

Pred. No.: 0.00197 Length: 681  
 Score: 12.00 Matches: 12  
 Percent Similarity: 100.00% Conservatave: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 12.37% Indels: 0  
 DB: 4 Gaps: 0

US-09-854-133-586 (1-97) x CFA279982 (1-681)

OY 23 ArgLysLysGluArgLysLysLysArgGluArgLys 34  
 |||||

Db 249 AGAAGAGAAAGAAAGAAAGAAAGAGAGAAG 284

RESULT 16

CH19F15314/c 41369 bp DNA linear PRI 22-MAR-1997

LOCUS CH19F15314 Homo sapiens DNA from chromosome 19p13.1 cosmid f15314, genomic

DEFINITION sequence.

ACCESSION AD000091

VERSION AD000091.1 GI:1905896  
 D19S11; chromosome 19; cytochrome P450; pseudogene.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE Identification of a CYP4F gene family at the D19s11 locus on 19p13.1

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 41369)

AUTHORS Lamerdin, J.E.

TITLE Direct Submission

COMMENT Submitted (07-NOV-1996) J.E. Lamerdin, Human Genome Center, Lawrence Livermore National Laboratory, 7000 East Ave, Livermore, CA, USA, 94551 jane@agtl.llnl.gov owetornak.llnl.gov  
 GSDB:S:1079875.  
 constructed at LLNL from flow-sorted chromosomes  
 from hybrid UV5HL9-5B, which carries chromosome 19 as its only human chromosome.

FEATURES

source

location/Qualifiers

1. .41369  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="19"  
 /clone="F15314"  
 /cell\_line="UV5HL9-5B"  
 /cell\_type="fibroblast"  
 /clone\_lib="L L19NC02 F2 chromosome 19-specific cosmid library"

complement(783. .1063)  
 /note="repeat match = HSAL06523; putative"  
 /rpt\_family="Alu"  
 complement(1029. .1072)  
 /note="repeat match = HSAL05178; putative"  
 /rpt\_family="Alu"

1396. .1464  
 /note="similarity: sp|Q08477; similarity to CPF3 (human); putative"  
 1401. .1461  
 /note="predicted exon, grail2exons\_human\_1.3; frame=2, forward strand, quality=marginal; putative"

1405. .1464  
 /note="similarity: sp|P33274; similarity to CP4F (rat); putative"  
 1408. .1461  
 /note="Human Leukotriene B4 omega-hydroxylase family pseudogene, exon 4; putative; does not fit consensus"

/number=4  
 /pseudo  
 1544. .1681  
 /note="similarity: sp|P33274; similarity to CP4F (rat); putative"

1544. .1681  
 /note="similarity: sp|Q08477; similarity to CPF3 (human); putative"

1553. .1680  
 /note="Human Leukotriene B4 omega-hydroxylase family pseudogene, exon 5; putative; does not fit consensus"

/number=5  
 /pseudo  
 1553. .1680  
 /note="predicted exon, grail2exons\_human\_1.3; frame=0, forward strand, quality=excellent; putative"

1985. .2244  
 /note="repeat match = HSAL00518; putative"  
 /rpt\_family="Alu"

2136. .2275  
 /note="repeat match = HSAL00299; putative"  
 /rpt\_family="Alu"

2395. .2416  
 /note="predicted exon, grail2exons\_human\_1.3; frame=0, forward strand, quality=excellent; putative"

misc\_feature

misc\_feature

misc\_feature

exon

misc\_feature

misc\_feature

repeat\_region

repeat\_region

repeat\_region

misc\_feature

```

misc_feature      forward strand, quality=good; putative"
                    2504..2586
repeat_region     /note="predicted exon, grail2exons_human_1.3; frame=2,
                    forward strand, quality=good; putative"
                    complement(2910..2977)
                    /note="repeat match = HSAL04183; putative"
                    /rpt_family="Alu"
                    complement(2924..3208)
exon              /note="repeat match = HSAL06258; putative"
                    /rpt_family="Alu"
                    3527..3648
                    /note="Human Leukotriene B4 omega-hydroxylase family
                    pseudogene, exon 6; putative; does not fit consensus"
                    /number=6
misc_feature      /pseudo
                    3527..3648
                    /note="predicted exon, grail2exons_human_1.3; frame=1,
                    forward strand, quality=excellent; putative"
                    3528..3650
misc_feature      /note="similarity: sp|P33274; similarity to CP4F (rat);
                    putative"
                    3528..3650
                    /note="similarity: sp|Q08477; similarity to CPF3 (human);
                    putative"
                    4378..4648
exon              /note="Human Leukotriene B4 omega-hydroxylase family
                    pseudogene, exon 7; putative; does not fit consensus"
                    /number=7
misc_feature      /pseudo
                    4380..4649
                    /note="similarity: sp|P33274; similarity to CP4F (rat);
                    putative"
                    4380..4649
misc_feature      /note="similarity: sp|Q08477; similarity to CPF3 (human);
                    putative"
                    4387..4648
misc_feature      /note="predicted exon, grail2exons_human_1.3; frame=1,
                    forward strand, quality=excellent; putative"
                    4684..4778
misc_feature      /note="predicted exon, grail2exons_human_1.3; frame=0,
                    forward strand, quality=marginal; putative"
                    4986..5052
exon              /note="Human Leukotriene B4 omega-hydroxylase family
                    pseudogene, exon 8; putative; does not fit consensus"
                    /number=8
misc_feature      /pseudo
                    4986..5052
                    /note="predicted exon, grail2exons_human_1.3; frame=2,
                    forward strand, quality=excellent; putative"
                    4987..5055
misc_feature      /note="similarity: sp|P33274; similarity to CP4F (rat);
                    putative"
                    4987..5055
misc_feature      /note="similarity: sp|Q08477; similarity to CPF3 (human);
                    putative"
                    5244..5375
misc_feature      /note="similarity: sp|P33274; similarity to CP4F (rat);
                    putative"
                    5244..5375
                    /note="similarity: sp|Q08477; similarity to CPF3 (human);
                    putative"
                    5244..5373
exon              /note="Human Leukotriene B4 omega-hydroxylase family
                    pseudogene, exon 9; putative; does not fit consensus"
                    /number=9
misc_feature      /pseudo
                    5244..5373
                    /note="predicted exon, grail2exons_human_1.3; frame=1,
                    forward strand, quality=excellent; putative"
                    complement(5596..5707)
                    /note="repeat match = HSAL05171; putative"
                    /rpt_family="Alu"
                    complement(5625..5924)
repeat_region

```

```

misc_feature      /note="repeat match = HSAL02250; putative"
                    /rpt_family="Alu"
                    6290..7080
repeat_region     /note="ss region (gap in top strand); putative"
                    complement(6665..6950)
                    /note="repeat match = HSAL00697; putative"
                    /rpt_family="Alu"
                    complement(6673..6879)
repeat_region     /note="repeat match = HSAL05100; putative"
                    /rpt_family="Alu"
                    complement(6884..6967)
repeat_region     /note="repeat match = HSAL06328; putative"
                    /rpt_family="Alu"
                    complement(7298..7445)
repeat_region     /note="repeat match = HSAL04736; putative"
                    /rpt_family="Alu"
                    complement(7331..7620)
repeat_region     /note="repeat match = Alu; putative"
                    /rpt_family="Alu"
                    complement(7548..7639)
repeat_region     /note="repeat match = HSAL02061; putative"
                    /rpt_family="Alu"
                    complement(7739..8057)
repeat_region     /note="repeat match = HSAL06256; putative"
                    /rpt_family="Alu"
                    complement(7837..8068)
repeat_region     /note="repeat match = HSAL00736; putative"
                    /rpt_family="Alu"
                    9050..9100
misc_feature      /note="ss region (gap in bottom strand); putative"
                    complement(9110..9354)
repeat_region     /note="repeat match = HSAL04630; putative"
                    /rpt_family="Alu"
                    complement(9173..9712)
repeat_region     /note="repeat match = HSAL01827; putative"
                    /rpt_family="Alu"
                    complement(9677..9723)
repeat_region     /note="repeat match = HSAL03456; putative"
                    /rpt_family="Alu"
                    complement(9818..9930)
repeat_region     /note="repeat match = HSAL00794; putative"
                    /rpt_family="Alu"
                    complement(9841..9977)
repeat_region     /note="repeat match = HSAL01882; putative"
                    /rpt_family="Alu"
                    complement(9873..9989)
repeat_region     /note="repeat match = HSAL05507; putative"
                    /rpt_family="Alu"

```

Alignment Scores:

Pred. No.:	Score:	Length:
0.07	12.00	41369
Percent Similarity:	100.00%	12
Best Local Similarity:	100.00%	0
Query Match:	12.37%	0
DB:	9	0

US-09-854-133-586 (1-97) x CH19P15314 (1-41369)

QY 23 ArgLysLysGluArgLysLysLysArgGluArgLys 34  
 DB 22456 AGAAGAAGAAAGAAAGAAAGAGAGAGAAAG 22421

RESULT 17  
 AC101290/c. 68527 bp DNA linear HTG 23-NOV-2001  
 LOCUS Mus musculus clone RP23-103D10, LOW-PASS SEQUENCE SAMPLING.  
 DEFINITION AC101290  
 ACCESSION AC101290.1 GI:17060065  
 VERSION HTG; HTGS\_PHASE0.  
 KEYWORDS Mus musculus.  
 SOURCE Mus musculus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
AUTHORS

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 68527)  
Bliren,B., Linton,L., Nusbaum,C. and Lander,E.  
Mus musculus, clone RP23-103D10

2 (bases 1 to 68527)  
Unpublished

Bliren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,  
Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B.,  
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,  
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,  
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,  
Ginde,S., Gorf,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,  
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,  
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,  
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,  
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNAL

## COMMENT

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L16367

Center clone name: 103\_D\_10

-----  
\* NOTE: This record contains 88 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 652: contig of 652 bp in length  
\* 653 752: gap of 100 bp  
\* 753 1436: contig of 684 bp in length  
\* 1437 1536: gap of 100 bp  
\* 1537 2226: contig of 690 bp in length  
\* 2227 2326: gap of 100 bp  
\* 2327 2993: contig of 667 bp in length  
\* 2994 3093: gap of 100 bp  
\* 3094 3785: contig of 692 bp in length  
\* 3786 3885: gap of 100 bp  
\* 3886 4575: contig of 690 bp in length  
\* 4576 4675: gap of 100 bp  
\* 4676 5365: contig of 690 bp in length  
\* 5366 5465: gap of 100 bp  
\* 5466 6163: contig of 698 bp in length  
\* 6164 6263: gap of 100 bp  
\* 6264 6962: contig of 699 bp in length  
\* 6963 7062: gap of 100 bp  
\* 7063 7741: contig of 679 bp in length

\* 7742 7841: gap of 100 bp  
\* 7842 8528: contig of 687 bp in length  
\* 8529 8628: gap of 100 bp  
\* 8629 9293: contig of 665 bp in length  
\* 9294 9393: gap of 100 bp  
\* 9394 10062: contig of 669 bp in length  
\* 10063 10162: gap of 100 bp  
\* 10163 10812: contig of 650 bp in length  
\* 10813 10912: gap of 100 bp  
\* 10913 11600: contig of 688 bp in length  
\* 11601 11700: gap of 100 bp  
\* 11701 12394: contig of 694 bp in length  
\* 12395 12494: gap of 100 bp  
\* 12495 13162: contig of 668 bp in length  
\* 13163 13262: gap of 100 bp  
\* 13263 13946: contig of 684 bp in length  
\* 13947 14046: gap of 100 bp  
\* 14047 14738: contig of 692 bp in length  
\* 14739 14838: gap of 100 bp  
\* 14839 15510: contig of 672 bp in length  
\* 15511 15610: gap of 100 bp  
\* 15611 16296: contig of 686 bp in length  
\* 16297 16396: gap of 100 bp  
\* 16397 17057: contig of 661 bp in length  
\* 17058 17157: gap of 100 bp  
\* 17158 17842: contig of 685 bp in length  
\* 17843 17942: gap of 100 bp  
\* 17943 18608: contig of 666 bp in length  
\* 18609 18708: gap of 100 bp  
\* 18709 19410: contig of 702 bp in length  
\* 19411 19510: gap of 100 bp  
\* 19511 20159: contig of 649 bp in length  
\* 20160 20259: gap of 100 bp  
\* 20260 20939: contig of 680 bp in length  
\* 20940 21039: gap of 100 bp  
\* 21040 21719: contig of 680 bp in length  
\* 21720 21819: gap of 100 bp  
\* 21820 22495: contig of 676 bp in length  
\* 22496 22595: gap of 100 bp  
\* 22596 23280: contig of 685 bp in length  
\* 23281 23380: gap of 100 bp  
\* 23381 24065: contig of 685 bp in length  
\* 24066 24165: gap of 100 bp  
\* 24166 24861: contig of 696 bp in length  
\* 24862 24961: gap of 100 bp  
\* 24962 25657: contig of 696 bp in length  
\* 25658 25757: gap of 100 bp  
\* 25758 26424: contig of 667 bp in length  
\* 26425 26524: gap of 100 bp  
\* 26525 27156: contig of 632 bp in length  
\* 27157 27256: gap of 100 bp  
\* 27257 27953: contig of 697 bp in length  
\* 27954 28053: gap of 100 bp  
\* 28054 28751: contig of 698 bp in length  
\* 28752 28851: gap of 100 bp  
\* 28852 29533: contig of 682 bp in length  
\* 29534 29633: gap of 100 bp  
\* 29634 30320: contig of 687 bp in length  
\* 30321 30420: gap of 100 bp  
\* 30421 31064: contig of 644 bp in length  
\* 31065 31164: gap of 100 bp  
\* 31165 31836: contig of 672 bp in length  
\* 31837 31936: gap of 100 bp  
\* 31937 32630: contig of 694 bp in length  
\* 32631 32730: gap of 100 bp  
\* 32731 33394: contig of 664 bp in length  
\* 33395 33494: gap of 100 bp  
\* 33495 34157: contig of 663 bp in length  
\* 34158 34257: gap of 100 bp  
\* 34258 34947: contig of 690 bp in length  
\* 34948 35047: gap of 100 bp  
\* 35048 35724: contig of 677 bp in length  
\* 35725 35824: gap of 100 bp



*	35825	36499:	contlg of 675 bp	in length
*	36500	36599:	gap of 100 bp	
*	37285	37284:	contlg of 685 bp	in length
*	37385	37384:	gap of 100 bp	
*	38179	38078:	contlg of 694 bp	in length
*	38179	38178:	gap of 100 bp	
*	38863	38862:	contlg of 684 bp	in length
*	38963	39652:	contlg of 690 bp	in length
*	39653	39752:	gap of 100 bp	
*	39753	40440:	contlg of 688 bp	in length
*	40441	40540:	gap of 100 bp	
*	40541	41236:	contlg of 696 bp	in length
*	41237	41336:	gap of 100 bp	
*	41337	42013:	contlg of 677 bp	in length
*	42014	42113:	gap of 100 bp	
*	42114	42762:	contlg of 649 bp	in length
*	42763	42862:	gap of 100 bp	
*	42863	43562:	contlg of 700 bp	in length
*	43563	43662:	gap of 100 bp	
*	43663	44367:	contlg of 705 bp	in length
*	44368	44467:	gap of 100 bp	
*	44468	45142:	contlg of 675 bp	in length
*	45143	45242:	gap of 100 bp	
*	45243	45915:	contlg of 673 bp	in length
*	45916	46015:	gap of 100 bp	
*	46016	46710:	contlg of 695 bp	in length
*	46711	46810:	gap of 100 bp	
*	46811	47491:	contlg of 681 bp	in length
*	47492	47591:	gap of 100 bp	
*	47592	48290:	contlg of 699 bp	in length
*	48291	48390:	gap of 100 bp	
*	48391	49063:	contlg of 673 bp	in length
*	49064	49163:	gap of 100 bp	
*	49164	49827:	contlg of 664 bp	in length
*	49828	49927:	gap of 100 bp	
*	49928	50597:	contlg of 670 bp	in length
*	50598	50697:	gap of 100 bp	
*	50698	51358:	contlg of 661 bp	in length
*	51359	51458:	gap of 100 bp	
*	51459	52141:	contlg of 683 bp	in length
*	52142	52241:	gap of 100 bp	
*	52242	52912:	contlg of 671 bp	in length
*	52913	53012:	gap of 100 bp	
*	53013	53689:	contlg of 677 bp	in length
*	53690	53789:	gap of 100 bp	

**Alignment Scores:**

```

Pred. No.:      0.108      Length:      685227
Score:          12.00      Matches:      12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    12.37% Indels: 0
DB:             2

```

US-09-854-133-586 (1-97) x AC101290 (1-68527)

Qy	23	ArgLysLysGluArgLysLysLysArgGluArgLys	34
Db	29955	AGAAAGAAAGAAAGAAAGAGAGAGAGAGAAA	29920

RESULT 18	
AC022364/C	
LOCUS	
DEFINITION	73557 bp DNA
Homo sapiens 12 BAC RP11-967K21 (Roswell Park Cancer Institute	linear
Human BAC Library) complete sequence.	PRI 23-MAR-2001
AC022364	
ACCESSION	

VERSION  
KEYWORDS  
SOURCE  
ORGANISM

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

## REFERENCE AUTHORS

REFERENCE	AUTHORS
1	(bases 1 to 73557)
	Muzny,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C., Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,S., Holloway,C., Hollins,B., Homsl,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louisedge,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N., Oyledo,R., Pace,A., Payton,B., Peety,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C., Shoochitari,N., Sisson,I., Sodergreen,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmant,K., Vasquez,L., Veta,V., Villalón,D., Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Kucherlapati,R. and Gibbs,R.

## JOURNAL REFERENCE

AUTHORS  
 TITLE  
 JOURNAL

## REFERENCE

**AUTHORS**  
**TITLE**  
**JOURNAL**

## REFERENCE

**AUTHORS**  
**TITLE**  
**JOURNAL**

**COMMENT**

On Jan 31, 2001 this sequence version replaced g1:12061318.  
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
[gc-helpe@bcm.tmc.edu](mailto:gc-helpe@bcm.tmc.edu)

**CLONE LENGTH:** This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.



## ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

## QUALSTAT-REPORT-----

## ----- Summary Statistics -----

Contig length: 73557  
Phrap values in estimate: 73121  
Average error rate (BCM-Phrap estimate): 0.000108988  
Fraction of Phrap values less than 40 : 0.00879364  
Number of consensus changing edits: 2  
Number of N's in consensus : 0

Position	Original+Context	Edited+Context
20701	tttttttt(n)gagacagagt	tttttttt(g)gagacagagt
37685	cttacacttg(g)ttttctccc	cttacacttg(t)ttttctccc

----- Distribution of Quality < 40 Bases -----

#	bases	5	10	15	20	25	30	35	40
5001									*
4501									*
4001									*
3501									*
3001									*
2501									*
2001									*
1501									*
1001									*
501									*
0		*	*	*	*	*	*	*	*

Phrap Value Range

Version: 1.01 gxf.

Location/Qualifiers

source

1. 73557

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="12"

/clone="RP11-967K21"

1. 199

misc\_feature /note="Overlaps bases 180995. 181193 of clone AC022079"

/function="Overlaps with adjacent clone AC022079"

41. 157  
/rpt\_family="T-rich"  
complement(219. 599)  
/rpt\_family="THEIC"  
1136. 1210  
/rpt\_family="MLT1K"  
complement(2086. 3012)  
/rpt\_family="L1PA16"  
4736. 4847  
/rpt\_family="L2"  
complement(4864. 5356)  
/rpt\_family="MLT1J"  
complement(5433. 5729)  
/rpt\_family="HERV16"  
complement(5730. 5810)  
/rpt\_family="HERS8A"  
complement(5937. 5955)  
/rpt\_family="HERV16"  
complement(5956. 6311)  
/rpt\_family="THEIC"  
complement(6312. 6479)  
/rpt\_family="HERV16"  
complement(6551. 7020)  
/rpt\_family="HERV16"  
complement(7018. 7465)  
/rpt\_family="HERV16"  
complement(7480. 7758)  
/rpt\_family="LTR16A"  
7897. 8232  
/rpt\_family="L2"  
8233. 8325  
/rpt\_family="L1M4"  
8326. 9024  
/rpt\_family="L1MC1"  
10770. 11148  
/rpt\_family="MSTA"  
complement(14187. 14760)  
/rpt\_family="L1MB3"  
complement(14761. 15060)  
/rpt\_family="AluY"  
complement(15061. 15688)  
/rpt\_family="L1MB3"  
complement(15689. 16119)

## Alignment Scores:

Pred. No.:	0.115	Length:	73557
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.37%	Indels:	0
DB:	9	Gaps:	0

US-09-854-133-586 (1-97) x AC022364 (1-73557)

QY 23 ArgLysLysGluArgLysLysLysArgGluArgLys 34

Db 37764 AGAAGAAAGAAAGAAAGAAAGAGAGAGAAAG 37729

## RESULT 19

AP000960 76163 bp DNA linear PRI 24-FEB-2000

LOCUS AP000960 Homo sapiens genomic DNA, chromosome 21q21.1-q21.2, clone:B800J1,

DEFINITION L156-APP region, complete sequence.

ACCESSION AP000960 GI:7077205

VERSION AP000960 HTG.

KEYWORDS Homo sapiens DNA, clone:B800J1.

SOURCE Homo sapiens DNA, clone:B800J1.

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 76163) Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

```
COMMENT      On Feb 24, 1980 this sequence version replaced gi:6997411  
FEATURES     location/Qualifiers  
source       1..76163
```

Alignment Scores:

Area: NO.:	0.119	Length:	76163
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.37%	Indels:	0
DB:	9		

US-09-854-133-586 (1-97) x AP000960 (1-76163)

QY 23 ArgLysLysGluArgLysLysArgGluArgLys 34  
 |||||  
 Db 19555 AGAAGAAAGAAAGAAAGAAAGAGAGAAAGGAAG 19590

RESULT 20	100167 bp	DNA	linear	PRI 24-FEB-2001
HSJ189G13/c				
LOCUS				
DEFINITION	Human DNA sequence from clone RPL-189G13 on chromosome 20. Contains			
	an RPL7A (60S ribosomal protein L7A) (SURF3) pseudogene, an RPS4			
	(40S ribosomal protein S4) pseudogene, ESTs, STSS and GSSs,			
	complete sequence.			
ACCESSION	AF131916			

ACCESSION	AL121916	Probe sequence.
VERSION	AL121916.14	
KEYWORDS	HTG; RPL7A; RPS4; SURF3	
SOURCE	human.	
ORGANISM	Homo sapiens	

**REFERENCE**  
AUTHORS  
Williams S  
1 (bases 1 to 100167)  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

**TITLE** Direct Submission  
**JOURNAL** Submitted (22-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire  
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk  
**COMMENT** requests: clonerequest@sanger.ac.uk  
On Apr 3, 2000 this sequence

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr20>  
**IMPORTANT:** This sequence is not the entire insert of clone RP1-189G13. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.  
 The true left end of clone RP1-189G13 is at 1 in this sequence. The true left end of clone RP5-1068H6 is at 100068 in this sequence. The true right end of clone RP5-1164C1 is at 13278 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 RP1-189G13 is from the library RPC1-1 constructed by the group of Pieter de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>  
**VECTOR:** pCYPAC2.

FEATURES	Location/Qualifiers
source	1..100167 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="20" /clone="RP1-189G13" /clone_lib="RPC1-1" 14..11225
repeat_region	/note="L1MA9 repeat: matches 4456..5781 of consensus" 1236..1327
misc_feature	/note="46 copies 2 mer to 80% conserved" complement(1276..1378)
repeat_region	/note="match: GSS: Em:AQ07965" 1376..1864
repeat_region	/note="L1MA9 repeat: matches 5783..6294 of consensus" 1987..2288
repeat_region	/note="AluX repeat: matches 1..310 of consensus" 2431..2702
repeat_region	/note="L2 repeat: matches 2241..2537 of consensus" 3194..3649
repeat_region	/note="L1PA13 repeat: matches 5702..6144 of consensus" 3779..3842
repeat_region	/note="32 copies 2 mer to 71% conserved" 4654..4751
repeat_region	/note="MIR repeat: matches 124..230 of consensus" 4788..5239
repeat_region	/note="L1MB7 repeat: matches 5481..5907 of consensus" 5240..5537
repeat_region	/note="AluJo repeat: matches 1..299 of consensus" 5538..5663
repeat_region	/note="L1MB7 repeat: matches 5363..5481 of consensus" 5664..5958
repeat_region	/note="AluX repeat: matches 1..297 of consensus" 5959..6048
repeat_region	/note="L1MB7 repeat: matches 5207..5363 of consensus" 6046..6269
misc_feature	/note="L1MB7 complement(6138..6621) /note="match: GSS: Em:AQ439122" 6539..8750
repeat_region	/note="L1MA9 repeat: matches 3474..5706 of consensus" 8751..9050
repeat_region	/note="AluX repeat: matches 1..300 of consensus" 9051..9362
repeat_region	/note="AluX repeat: matches 1..312 of consensus" 9363..9762
repeat_region	/note="L1MA9 repeat: matches 5706..6064 of consensus" 9763..9800
misc_feature	/note="L1R16C repeat: matches 350..387 of consensus" complement(9941..10370)
misc_feature	/note="match: GSS: Em:AQ696486" 10332..10825
repeat_region	/note="match: GSS: Em:AQ672010" 11276..11654
	/note="L1R16C repeat: matches 1..366 of consensus"

```

repeat_region      11928..12124
                    /note="HERV repeat: matches 4464..4658 of consensus"
repeat_region      12503..13301
                    /note="HERV repeat: matches 3127..3940 of consensus"
misc_feature       complement(13588..13691)
                    /note="match: GSS: Em:AQ548735 Em:AQ800523
                    match: STS: Em:AQ548735"
repeat_region      13724..13949
                    /note="LIM4 repeat: matches 3331..3553 of consensus"
repeat_region      13951..14855
                    /note="LIM4 repeat: matches 5440..6317 of consensus"
repeat_region      14836..16206
                    /note="LIM4 repeat: matches 2655..4042 of consensus"
repeat_region      16207..16426
                    /note="LIM4 repeat: matches 1..230 of consensus"
repeat_region      16427..16523
                    /note="LIM4 repeat: matches 2560..2655 of consensus"
repeat_region      16531..16740
                    /note="LIM4 repeat: matches 353..558 of consensus"
repeat_region      16746..16805
                    /note="LIM4 repeat: matches 2 mer tg 90% conserved"
repeat_region      16806..17085
                    /note="LIM4 repeat: matches 58..345 of consensus"
repeat_region      17102..17182
                    /note="LIM4 repeat: matches 1..88 of consensus"
repeat_region      17192..17323
                    /note="LIM4 repeat: matches 3195..3319 of consensus"
repeat_region      17579..17752
                    /note="LIM4 repeat: matches 6..186 of consensus"
misc_feature       17998..18371
                    /note="match: GSS: Em:B99526"
repeat_region      18871..19048
                    /note="LIM4 repeat: matches 85..262 of consensus"
repeat_region      19049..19554
                    /note="LIM4 repeat: matches 19..547 of consensus"
repeat_region      19558..19581
                    /note="LIM4 repeat: matches 12 copies 2 mer tt 100% conserved"
repeat_region      20668..20913
                    /note="LIM4 repeat: matches 2 mer ct 81% conserved"
repeat_region      21007..21238
                    /note="LIM4 repeat: matches 2..229 of consensus"
misc_feature       21496..21915
                    /note="match: GSS: Em:AQ20755"
misc_feature       21510..21716
                    /note="match: GSS: Em:AQ297835"
repeat_region      21582..21657
                    /note="LIM4 repeat: matches 67..144 of consensus"
repeat_region      21808..21901
                    /note="LIM4 repeat: matches 254..349 of consensus"
misc_feature       complement(21830..22310)
                    /note="match: GSS: Em:AQ141400"
repeat_region      22351..23665
                    /note="LIM4 repeat: matches 2384..3718 of consensus"
repeat_region      26991..27304
                    /note="LIM4 repeat: matches 2920..3242 of consensus"
repeat_region      27392..27480
                    /note="LIM4 repeat: matches 105..189 of consensus"
repeat_region      27655..27840
                    /note="LIM4 repeat: matches 71..262 of consensus"
repeat_region      28167..30593
                    /note="LIM4 repeat: matches 4017..5403 of consensus"
repeat_region      30621..30679
                    /note="LIM4 repeat: matches 6853..6918 of consensus"
repeat_region      30685..31400
                    /note="LIM4 repeat: matches 5504..6224 of consensus"
repeat_region      31431..31557
                    /note="LIM4 repeat: matches 1..130 of consensus"
repeat_region      31950..31984
                    /note="LIM4 repeat: matches 513..547 of consensus"
repeat_region      31980..32672
                    /note="LIM4 repeat: matches 96..773 of consensus"
misc_feature       32902..33244
                    /note="match: GSS: Em:B32572"

```

```

repeat_region      34122..34247
                    /note="L2 repeat: matches 2572..2710 of consensus"
misc_feature       complement(34327..34694)
                    /note="match: GSS: Em:AQ120726"
misc_feature       34676..35116
                    /note="match: GSS: Em:B16156"
repeat_region      34720..35038
                    /note="LIM4 repeat: matches 226..547 of consensus"
repeat_region      35056..35397
                    /note="LIM4 repeat: matches 240..583 of
                    consensus"
repeat_region      35650..35843
                    /note="LIM4 repeat: matches 688..890 of
                    consensus"
repeat_region      36180..36250
                    /note="LIM4 repeat: matches 1108..1177 of
                    consensus"
repeat_region      36329..36451
                    /note="LIM4 repeat: matches 6046..6172 of consensus"
repeat_region      36445..36586
                    /note="LIM4 repeat: matches 5345..5484 of consensus"

```

Alignment Scores:

Pred. No.:	0.151	Length:	100167
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.37%	Indels:	0
DB:	9	Gaps:	0

US-09-854-133-586 (1-97) x HSC189G13 (1-100167)

QY 23 ArglyslsGluarglyslsArgGluArglysls 34  
 Db 20789 AGAAGAAAGAAAGAAAGAAAGAAAGAAAG 20754

RESULT 21  
 AC090033/c 125267 bp DNA linear MAM 21-JUL-2001  
 LOCUS  
 DEFINITION Felis catus clone RP86-523H23, complete sequence.  
 AC090033  
 AC090033.2 GI:14993725  
 VERSION  
 KEYWORDS HTG.  
 SOURCE Felis catus.  
 ORGANISM Felis catus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.

REFERENCE  
 AUTHORS  
 1 (bases 1 to 125267)  
 Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,  
 Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L.,  
 Granite, S., Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E.,  
 Lee-Lin, S.-O., Legaspi, R., Lim, M., Maduro, Q.L., Maduro, V.B.,  
 Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J.,  
 Pearson, R., Prasad, A., Shevchenko, Y., Snyder, B., Stantrifop, S.,  
 Thomas, J.W., Thomas, P.J., Touchman, J.W., Tsugeon, C., Vogt, J.L.,  
 Walker, M.A., Wetherby, K.D., Zhang, L.-H. and Green, E.D.  
 NISC Comparative Sequencing Initiative

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 2 (bases 1 to 125267)  
 Green, E.D.  
 Direct Submission  
 Submitted (10-FEB-2001) NIH Intramural Sequencing Center, 8717  
 Grovemont Circle, Gaithersburg, MD 20877, USA  
 3 (bases 1 to 125267)  
 Green, E.D.  
 Direct Submission  
 Submitted (21-JUL-2001) NIH Intramural Sequencing Center, 8717  
 Grovemont Circle, Gaithersburg, MD 20877, USA

COMMENT  
 ----- Genome Center  
 On Jul 21, 2001 this sequence version replaced gi:12739798.  
 Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: <http://www.nisc.nih.gov>

Contact: nisc\_mouse@hgrl.nih.gov  
----- Project Information  
Center project name: awy  
Center clone name: 523H23

This sequence was finished as follows unless otherwise noted:  
all regions were double-stranded, sequenced with an  
alternate chemistry, or covered by high quality data  
(i.e., phred quality >= 30); an attempt was made to resolve  
all sequencing problems, such as compressions and repeats;  
all regions were covered by at least one plasmid subclone  
or more than one M13 subclone; and the assembly was confirmed  
by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of  
this clone unless otherwise noted. If there are overlapping  
clones, the overlaps are noted in the beginning and end of  
the Features section.

FEATURES  
Source

Location/Qualifiers  
1. .125267  
/organism="Felis catus"  
/db\_xref="taxon:9685"  
/clone="RP86-523H23"  
/clone\_lib="RP86"  
1. .31338

/note="clone overlaps with GenBank Accession Number  
AC087731 (nucleotides 94316-125661) clone RP86-45908  
(center project name awx)"  
50020. .50113  
misc\_feature /note="single clone coverage"

BASE COUNT 37454 a 22462 c 24472 g 40879 t  
ORIGIN

Alignment Scores:

Pred. No.: 0.183 Length: 125267  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
DB: 4 Gaps: 0

US-09-854-133-586 (1-97) x AC090033 (1-125267)

QY 23 ArgLysLysGluArgLysLysLysArgGluArgLys 34  
|||||  
DB 548 AGGAAGAAAGAGAGAAAGAAAGAAAGAGAAAG 513  
|||||

RESULT 22

AC087731/c

LOCUS AC087731 125661 bp DNA linear MAM 03-NOV-2001  
DEFINITION Felis catus clone RP86-45908, complete sequence.  
AC087731  
AC087731.2 GI:12739794  
VERSION  
KEYWORDS HTG.  
SOURCE Felis catus.  
ORGANISM Felis catus.

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
1 (bases 1 to 125661)  
Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,  
Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S.,  
Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E., Lee-Jin, S.-Q.,  
Legaspi, R., Lim, M., Maduro, Q.L., Maduro, V.B., Mastello, C.,  
Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Prasad, A.,  
Shevchenko, Y., Snyder, B., Stantrop, S., Thomas, J.W., Thomas, P.J.,  
Tiongson, E.E., Touchman, J.W., Tsurgenev, C., Vogt, J.L., Walker, M.A.,  
Wetherby, K.D., Zhang, L.-H. and Green, E.D.  
NISC Comparative Sequencing Initiative  
Unpublished  
2 (bases 1 to 125661)  
Green, E.D.  
Direct Submission  
Submitted (19-JAN-2001) NIH Intramural Sequencing Center, 8717

REFERENCE  
AUTHORS Grovmont Circle, Gaithersburg, MD 20877, USA  
3 (bases 1 to 125661)  
Green, E.D.

TITLE Direct Submission

JOURNAL Submitted (10-FEB-2001) NIH Intramural Sequencing Center, 8717  
Grovmont Circle, Gaithersburg, MD 20877, USA  
4 (bases 1 to 125661)  
Green, E.D.

REFERENCE  
AUTHORS Direct Submission

TITLE Submitted (03-NOV-2001) NIH Intramural Sequencing Center, 8717  
JOURNAL Grovmont Circle, Gaithersburg, MD 20877, USA  
On Feb 10, 2001 this sequence version replaced g1:12313759.

COMMENT ----- Genome Center

Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: http://www.nisc.nih.gov  
Contact: nisc\_mouse@hgrl.nih.gov  
----- Project Information  
Center project name: awx  
Center clone name: 459008

This sequence was finished as follows unless otherwise noted:  
all regions were double-stranded, sequenced with an  
alternate chemistry, or covered by high quality data  
(i.e., phred quality >= 30); an attempt was made to resolve  
all sequencing problems, such as compressions and repeats;  
all regions were covered by at least one plasmid subclone  
or more than one M13 subclone; and the assembly was confirmed  
by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of  
this clone unless otherwise noted. If there are overlapping  
clones, the overlaps are noted in the beginning and end of  
the Features section.

FEATURES  
Source

Location/Qualifiers  
1. .125661  
/organism="Felis catus"  
/db\_xref="taxon:9685"  
/clone="RP86-45908"  
/clone\_lib="RP86"  
1. .43  
misc\_feature /note="clone has a very small overlap with GenBank  
Accession Number AC087861 (nucleotides 107321-107365)  
clone RP86-49M22 (center project name awx)"  
47870. .47882

misc\_feature /note="single clone coverage"  
49741. .49820  
/note="single clone coverage"  
50031. .50069  
/note="single clone coverage"  
67131. .67180  
/note="single clone coverage"  
94316. .125661  
/note="clone overlaps with GenBank Accession Number  
AC090033 (nucleotides 1-31338) clone RP86-523H23 (center  
project name awx)"

BASE COUNT 37714 a 24480 c 24041 g 39426 t  
ORIGIN

Alignment Scores:

Pred. No.: 0.184 Length: 125661  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
DB: 4 Gaps: 0

US-09-854-133-586 (1-97) x AC087731 (1-125661)

QY 23 ArgLysLysGluArgLysLysLysArgGluArgLys 34  
|||||  
DB 94863 AGGAAGAAAGAGAGAAAGAAAGAAAGAGAAAG 94828  
|||||



RESULT 23	AC105404/c	127475 bp	DNA	linear	HTG 21-MAY-2002
LOCUS	AC105404				
DEFINITION	Mus musculus chromosome UNK clone RP24-112M13, WORKING DRAFT SEQUENCE, 6 unordered pieces.				
ACCESSION	AC105404				
VERSION	AC105404.3	GI:21039970			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.				
SOURCE	Mus musculus.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 127475)				
TITLE	McPherson, J.D. and Waterston, R.H.				
JOURNAL	The sequence of Mus musculus clone				
REFERENCE	Unpublished				
AUTHORS	2 (bases 1 to 127475)				
TITLE	McPherson, J.D. and Waterston, R.H.				
JOURNAL	Direct Submission				
REFERENCE	Submitted (03-JAN-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA				
AUTHORS	3 (bases 1 to 127475)				
TITLE	McPherson, J.D. and Waterston, R.H.				
JOURNAL	Direct Submission				
REFERENCE	Submitted (21-MAY-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA				
COMMENT	On May 21, 2002 this sequence version replaced gi:18767639.				

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_BB0112M13
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 127786 bases at least Q40
Consensus quality: 128075 bases at least Q30
Consensus quality: 128255 bases at least Q20
Insert size: 13200; agarose-ff
Insert size: 129273; sum-of-contigs
Quality coverage: 14.01 in Q20 bases; agarose-ff
Quality coverage: 12.18 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 2153: contig of 2153 bp in length
*
* 2154 2253: gap of unknown length
*
* 2254 18501: contig of 16248 bp in length
*
* 18502 18601: gap of unknown length
*
* 18602 40805: contig of 22204 bp in length
*
* 40806 40905: gap of unknown length
*
* 40906 81368: contig of 40463 bp in length
*
* 81369 81468: gap of unknown length
*
* 81469 127159: contig of 45691 bp in length
*
* 127160 127259: gap of unknown length
*
* 127260 127475: contig of 216 bp in length.
*
FEATURES
source
1..127475
/organism="Mus musculus"
/db_xref="taxon:10090"

```

```

/chromosome="UNK"
/clone="RP24-112M13"
1. .2153
/note="assembly_name:Contig31"
2254. .18501
/note="assembly_name:Contig32"
18602. .40805
/note="assembly_name:Contig33"
40906. .81368
/note="assembly_name:Contig34"
81469. .127159
/note="assembly_name:Contig35"
127260. .127475
/note="assembly_name:Contig11"
BASE COUNT    41143 a    23490 c    23116 g    39205 t
ORIGIN
Alignment Scores:
Pred. No.:      0.186          length:      127475
Score:          12.00         Matches:      12
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:     12.37%      Indels:        0
DB:              Gaps:        2
US-09-854-133-586 (1-97) x AC105404 (1-127475)
QY      23  ARGLYSLYGLUARGLYSLYSLYsARGLUArgLYs  34
|||||
1816  AGAAGAAGAAAGAAAGAAAGAAAGAGAGAAAGAAAG  1781

```

RESULT 24				
AC099249/c				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
AC099249	143786 bp	DNA	linear	HTG 12-JUL-2002
Rattus norvegicus clone CH230-191L15,	***, 84 unordered pieces.			*** SEQUENCING IN PROGRESS
AC099249				
AC099249.5	G1:21729805			
HTG; HTGS_PHASE1.				
Norway rat.				
Rattus norvegicus				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;				
Rattus.				
1 (bases 1 to 143786)				
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaraltinge,H.C., Are,J.R., Ayele,M., Banks,T., Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,R.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Doutheite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsf., Howard,S., Huber,J., Huliyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivest,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,C., Liu,J., Liu,W., Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newtonson,J., Newtonson,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokekwu,S., Oguh,M., Okwuonu,G.,				

Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
Peters, L., Pickens, R., Primus, E., Pu, L.L., Qules, M., Ren, Y.,  
Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,  
Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,  
Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
Weinstock, G. and Gibbs, R.

Unpublished  
2 (bases 1 to 143786)  
Worley, K.C.

Direct Submission  
Submitted (09-NOV-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 143786)  
Worley, K.C.

Direct Submission  
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

On Jul 11, 2002 this sequence version replaced gi:18846141.

## COMMENT

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

Project Information

Center project name: GUS

Center clone name: CH230-191L15

Summary Statistics

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 49247 bases at least Q40

Consensus quality: 52098 bases at least Q30

Consensus quality: 53978 bases at least Q20

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 84 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1284: contig of 1284 bp in length  
\* 1285 1384: gap of unknown length  
\* 1385 2582: contig of 1198 bp in length  
\* 2583 2682: gap of unknown length  
\* 2683 3710: contig of 1028 bp in length  
\* 3711 3810: gap of unknown length  
\* 3811 5432: contig of 1622 bp in length  
\* 5433 5532: gap of unknown length  
\* 5533 6996: contig of 1464 bp in length  
\* 6997 7096: gap of unknown length  
\* 7097 8254: contig of 1158 bp in length  
\* 8255 8354: gap of unknown length  
\* 8355 9457: contig of 1103 bp in length  
\* 9458 9557: gap of unknown length  
\* 9558 10583: contig of 1026 bp in length  
\* 10584 10683: gap of unknown length  
\* 10684 11703: contig of 1020 bp in length  
\* 11704 11803: gap of unknown length  
\* 11804 12807: contig of 1004 bp in length  
\* 12808 12907: gap of unknown length

12908 13913: contig of 1006 bp in length  
\* 13914 14013: gap of unknown length  
\* 14014 15081: contig of 1068 bp in length  
\* 15082 15181: gap of unknown length  
\* 15182 16764: contig of 1583 bp in length  
\* 16765 16864: gap of unknown length  
\* 16865 17882: contig of 1018 bp in length  
\* 17883 17982: gap of unknown length  
\* 17983 19051: contig of 1069 bp in length  
\* 19052 19151: gap of unknown length  
\* 19152 20790: contig of 1639 bp in length  
\* 20791 20890: gap of unknown length  
\* 20891 21907: contig of 1017 bp in length  
\* 21908 22007: gap of unknown length  
\* 22008 23050: contig of 1043 bp in length  
\* 23051 23150: gap of unknown length  
\* 23151 24759: contig of 1609 bp in length  
\* 24760 24859: gap of unknown length  
\* 24860 26246: contig of 1387 bp in length  
\* 26247 26347: gap of unknown length  
\* 26348 27642: contig of 1296 bp in length  
\* 27643 27742: gap of unknown length  
\* 27743 29353: contig of 1611 bp in length  
\* 29354 29454: gap of unknown length  
\* 29455 30931: contig of 1478 bp in length  
\* 30932 31031: gap of unknown length  
\* 31032 32100: contig of 1069 bp in length  
\* 32101 32200: gap of unknown length  
\* 32201 34054: contig of 1854 bp in length  
\* 34055 34154: gap of unknown length  
\* 34155 35776: contig of 1622 bp in length  
\* 35777 35876: gap of unknown length  
\* 35877 37330: contig of 1454 bp in length  
\* 37331 37430: gap of unknown length  
\* 37431 38957: gap of unknown length  
\* 38958 39058: gap of unknown length  
\* 39059 40652: contig of 1595 bp in length  
\* 40653 40752: gap of unknown length  
\* 40753 42109: contig of 1357 bp in length  
\* 42110 42210: gap of unknown length  
\* 42211 44088: contig of 1879 bp in length  
\* 44089 44188: gap of unknown length  
\* 44189 45516: contig of 1328 bp in length  
\* 45517 45616: gap of unknown length  
\* 45617 47045: contig of 1429 bp in length  
\* 47046 47145: gap of unknown length  
\* 47146 48976: contig of 1831 bp in length  
\* 48977 49076: gap of unknown length  
\* 49077 50820: contig of 1744 bp in length  
\* 50821 50920: gap of unknown length  
\* 50921 52285: contig of 1365 bp in length  
\* 52286 52385: gap of unknown length  
\* 52386 53450: contig of 1065 bp in length  
\* 53451 53550: gap of unknown length  
\* 53551 54862: contig of 1312 bp in length  
\* 54863 54962: gap of unknown length  
\* 54963 56200: contig of 1238 bp in length  
\* 56201 56300: gap of unknown length  
\* 56301 57821: contig of 1521 bp in length  
\* 57822 57921: gap of unknown length  
\* 57922 58928: contig of 1007 bp in length  
\* 58929 59028: gap of unknown length  
\* 59029 60205: contig of 1177 bp in length  
\* 60206 60305: gap of unknown length  
\* 60306 61533: contig of 1228 bp in length  
\* 61534 61633: gap of unknown length  
\* 61634 62809: contig of 1176 bp in length  
\* 62810 62909: gap of unknown length  
\* 62910 63960: contig of 1051 bp in length  
\* 63961 64060: gap of unknown length  
\* 64061 65563: contig of 1503 bp in length  
\* 65564 65663: gap of unknown length  
\* 65664 67365: contig of 1702 bp in length

[REDACTED]



	misc_feature	25546.	.30628	/note="assembly-fragment"	
	misc_feature	30729.	.37323	/note="assembly-fragment"	
	misc_feature	37424.	.42575	/note="assembly-fragment"	
	misc_feature	42676.	.79708	/note="assembly-fragment"	
	misc_feature	79809.	.86724	/note="assembly-fragment"	
	misc_feature	86825.	.109636	/note="assembly-fragment"	
	misc_feature	109737.	.139954	/note="assembly-fragment"	
	misc_feature	140055.	.145722	/note="assembly-fragment"	
	misc_feature	clone_end:	T7		
COUNT		vector_side:right"			
NN	44952	a	27253	c	26791 g
			45325	t	
			1401	others	

Alignment Scores:	
Pred. No.:	0.209
Score:	12.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	12.37%
DB:	2
Length:	145722
Matches:	12
Conservative:	0
Mismatch:	0
Indels:	0
Gaps:	0

US-09-854-133-586 (1-97) x AC015503 (1-145722)

QY 23 ArgLysLysGLuArgLysLysArgGLuArgLys 34  
 Db 104218 AGAAGAGAAAGAGAAAAAGAAAGAGAAAGAAAG 104253

RESULT	26
AC004932/C	
LOCUS	AC004932
DEFINITION	Homo sapiens clone RP5-943F2, complete sequence.
ACCESSION	AC004932
VERSION	AC004932.4
KEYWORDS	GI:13446338
SOURCE	HTG.
ORGANISM	Homo sapiens.
	Homo sapiens
	Homo sapiens

**REFERENCE**  
1 (bases 1 to 151183)  
**AUTHORS**

JOURNAL REFERENCE AUTHORS TITLE JOURNAL	Unpublished 2 (bases 1 to 151183) Waterston, R.H. Direct Submission Submitted (12-JUN-1998) Genome

REFERENCE	3 (bases 1 to 151183)
AUTHORS	Waterston, R. H.
TITLE	Direct Submission
JOURNAL	Submitted (07-JUL-2000) Genome

REFERENCE	4 (bases 1 to 151183)	
AUTHORS	Waterston, R.	
TITLE	Direct Submission	
JOURNAL	Submitted (30-SEP-2000)	
REFERENCE	University, 4444 Forest	Department of Genetics, Washington
AUTHORS	5 (bases 1 to 151183)	
TITLE	Waterston, R. H.	
JOURNAL	Direct Submission	
	Submitted (25-MAR-2001)	Genome Sequencing Project

University School of Medicine, Washington  
MO 63108, USA

```
COMMENT      On Mar 25, 2001 this sequence version replaced gl:8954173
FEATURES
    Center    project name: H_DJ0943F02.
    Location/Qualifiers
        source          1..151183
```

BASE COUNT	ORIGIN
42404 a	34087 c 33409 g 41283 t

Alignment Scores:	
Pred. No.:	0.216
Score:	12.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	12.37%
DB:	9
Length:	151183
Matches:	12
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-854-133-586 (1-97) x AC004932 (1-151183)

QY	23	ArgIysLysGluArgIysLysLysArgGluArgLys	34
Db	31832	AGAAAAAGAAAGAAAAAGAAAGACAGAAAGAAAA	31797

RESULT	27
AC011275	
LOCUS	
DEFINITION	AC011275 Homo sapiens clone RP11-13J15, WORKING DRAFT SEQUENCE, 7 unordered pieces.
ACCESSION	AC011275
VERSION	AC011275.3 GI:10198431
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	Homo sapiens.

ORGANISM	REFERENCE
Homo sapiens	
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
1 (bases 1 to 152959)	
Birren, B., Linton, L., Nusbaum, C. and Lander, E.	
Homo sapiens, clone RP11-13J15	
Unpublished	
2 (bases 1 to 152959)	
Birren, B., Linton, L., Nusbaum, C. and Lander, E.	

Baldwin, J., Barna, N., Beckerly, R., Lander, E., Allen, N., Anderson, M., Brown, A., Castle, A., Colangelo, M., Boguslavsky, L., Boukhgalter, B., Cooke, P., Dearellano, K., Dewar, K., Collins, S., Collymore, A., Ferreira, P., FitzHugh, W., Forrest, C., Domino, M., Donelan, L., Doyle, M., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehoczký, J., Lieu, C., Locke, K., MacDonald, P., Marguis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

**TITLE**  
**JOURNAL**  
 Submitted (05-OCT-1999) Whiteboard

COMMENT  
On Sep 20, 2000 this sequence version replaced gl:9930771.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/PW/Repeats/>

Genome Center  
Center for

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

Project Information  
Center project name: I.3338

Center clone name: 13\_J\_15

Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads



Assembly program: Phrap; version 0.960731  
Consensus quality: 147891 bases at least Q40  
Consensus quality: 150486 bases at least Q30  
Consensus quality: 151334 bases at least Q20  
Insert size: 157000; agarose-fp  
Insert size: 152359; sum-of-contigs  
Quality coverage: 5.1 in Q20 bases; agarose-fp  
Quality coverage: 5.3 in Q20 bases; sum-of-contigs

\*\*\*\*\*  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 7 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*\*\*\*\*

1 25856: contig of 25856 bp in length  
\* 25857 25956: gap of 100 bp  
\* 25957 29722: contig of 3766 bp in length  
\* 29723 29822: gap of 100 bp  
\* 29823 34740: contig of 4918 bp in length  
\* 34741 34840: gap of 100 bp  
\* 34841 50346: contig of 15506 bp in length  
\* 50347 50446: gap of 100 bp  
\* 50447 115328: contig of 64882 bp in length  
\* 115329 115428: gap of 100 bp  
\* 115429 137085: contig of 21657 bp in length  
\* 137086 137185: gap of 100 bp  
\* 137186 152959: contig of 15774 bp in length.

FEATURES  
source  
1. .152959  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-13J15"  
/clone\_lib="RPC1-11 Human Male BAC"  
1. .25856  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left"

misc\_feature  
1. .25856  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left"

misc\_feature  
25957. .29722  
/note="assembly\_fragment"

misc\_feature  
29823. .34740  
/note="assembly\_fragment"

misc\_feature  
34841. .50346  
/note="assembly\_fragment"

misc\_feature  
50447. .115328  
/note="assembly\_fragment"

misc\_feature  
115429. .137085  
/note="assembly\_fragment"

misc\_feature  
137186. .152959  
/note="assembly\_fragment"

clone\_end:T7  
vector\_side:right"

BASE COUNT 48218 a 28547 c 28771 g 46796 t 627 others  
ORIGIN

#### Alignment Scores:

Pred. No.: 0.218 Length: 152959  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
DB: 2 Gaps: 0

US-09-854-133-586 (1-97) x AC011275 (1-152959)

OY 23 ArglyslsGluArglyslsArgGluArglys 34  
|||||

Db 53242 AGAAGAAAGAAAGAAAGAAAGAGAGAAAGAGAG 53277

RESULT 28  
AC011492/c

LOCUS AC011492 153064 bp DNA linear PRI 25-MAY-2002  
DEFINITION Homo sapiens chromosome 19 clone CTB-187L3, complete sequence.  
AC011492  
AC011492.8 GI:21206230  
HTG.  
Homo sapiens.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
DOI Joint Genome Institute and Stanford Human Genome Center.  
1 (bases 1 to 153064)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
Direct Submission  
Unpublished  
2 (bases 1 to 153064)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 153064)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
Direct Submission  
Submitted (25-MAY-2002) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
On May 25, 2002 this sequence version replaced gi:14971177.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 100% of Sequence;  
Estimated Total Number of Errors is 0.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
DOE Joint Genome Institute and Stanford Human Genome Center.  
Submitted (25-MAY-2002) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
On May 25, 2002 this sequence version replaced gi:14971177.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 100% of Sequence;  
Estimated Total Number of Errors is 0.

FEATURES  
source  
1. .153064  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="CTB-187L3"

BASE COUNT 35555 a 38286 c 37803 g 41420 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.218 Length: 153064  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x AC011492 (1-153064)

OY 23 ArglyslsGluArglyslsArgGluArglys 34  
|||||

Db 150690 AGAAGAAAGAAAGAAAGAAAGAGAGAGAGAAAG 150655

#### RESULT 29

AC022290/c 156795 bp DNA linear HTG 09-SEP-2000  
LOCUS Homo sapiens clone RP11-12L19, WORKING DRAFT SEQUENCE, 25 unordered  
DEFINITION

AC022290  
AC022290.3 GI:10045347  
HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
Homo sapiens.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 156795)  
Homo sapiens chromosome, clone RP11-12L19  
Unpublished  
2 (bases 1 to 156795)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,  
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Unpublished  
2 (bases 1 to 156795)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,  
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,

TITLE  
JOURNAL

## COMMENT

Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,  
Dearellano, K., Dewar, K., Domino, M., Doyle, M., Fenestor, J.,  
Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J.,  
Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K.,  
Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,  
McPheeters, R., Meldrim, J., Menus, L., Morrow, J., Naylor, J.,  
Norman, C.H., O'Connor, T., O'Donnell, P., Olivari, T.M., Peterson, K.,  
Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,  
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,  
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W.J.,  
Zimmer, A. and Zody, M.

Submitted (28-JAN-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Sep 9, 2000 this sequence version replaced gi:6850471.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

Project Information  
Center project name: L3293

Center clone name: 12\_L19

Summary Statistics  
Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 14511 bases at least Q40  
Consensus quality: 150979 bases at least Q30  
Insert size: 17000; agarose-fp  
Insert size: 154395; sum-of-contigs  
Quality coverage: 4.2 in Q20 bases; agarose-fp  
Quality coverage: 4.6 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently  
consists of 25 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 8317: contig of 8317 bp in length  
8318 8417: gap of 100 bp  
8418 9479: contig of 1062 bp in length  
9480 9579: gap of 100 bp  
9580 10818: contig of 1239 bp in length  
10819 10918: gap of 100 bp  
10919 12456: contig of 1538 bp in length  
12457 12556: gap of 100 bp  
12557 14941: contig of 2385 bp in length  
14942 15041: gap of 100 bp  
15042 17304: contig of 2263 bp in length  
17305 17404: gap of 100 bp  
17405 19807: contig of 2403 bp in length  
19808 19907: gap of 100 bp  
19908 22367: contig of 2460 bp in length  
22368 22467: gap of 100 bp  
22468 25490: contig of 3023 bp in length  
25491 25590: gap of 100 bp  
25591 29482: contig of 3892 bp in length  
29483 29582: gap of 100 bp  
29583 32815: contig of 3233 bp in length  
32816 32915: gap of 100 bp  
32916 36024: contig of 3109 bp in length  
36025 36124: gap of 100 bp

FEATURES  
source

36125 41302: contig of 5178 bp in length  
41303 41402: gap of 100 bp  
41403 45395: contig of 3993 bp in length  
45396 45495: gap of 100 bp  
45496 49599: contig of 4104 bp in length  
49600 49699: gap of 100 bp  
49700 68482: contig of 18783 bp in length  
68483 68582: gap of 100 bp  
68583 75750: contig of 7168 bp in length  
75751 75850: gap of 100 bp  
75851 81623: contig of 5773 bp in length  
81624 81723: gap of 100 bp  
81724 90287: contig of 8564 bp in length  
90288 90387: gap of 100 bp  
90388 99686: contig of 9299 bp in length  
99687 99786: gap of 100 bp  
99787 110519: contig of 10733 bp in length  
110520 110619: gap of 100 bp  
110620 123065: contig of 12446 bp in length  
123066 123165: gap of 100 bp  
123166 138417: contig of 15252 bp in length  
138418 138517: gap of 100 bp  
138518 151810: contig of 13293 bp in length  
151811 151910: gap of 100 bp  
151911 156795: contig of 4885 bp in length.

Location/Qualifiers

1. 156795

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="RP11-12L19"

/clone\_lib="RPCT-11 Human Male BAC"

1. 8317

/note="assembly-fragment"

clone\_end:sp6

vector\_side:left

8418. 9479

/note="assembly-fragment"

9580. 10818

/note="assembly-fragment"

10919. 12456

/note="assembly-fragment"

12557. 14941

/note="assembly-fragment"

15042. 17304

/note="assembly-fragment"

17405. 19807

/note="assembly-fragment"

19908. 22367

/note="assembly-fragment"

22468. 25490

/note="assembly-fragment"

25591. 29482

/note="assembly-fragment"

29583. 32815

/note="assembly-fragment"

32916. 36024

/note="assembly-fragment"

36125. 41302

/note="assembly-fragment"

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

misc\_feature

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

/note="assembly-fragment"

```
misc_feature      99787..110519
                  /note="assembly_fragment"
misc_feature      110620..123065
                  /note="assembly_fragment"
misc_feature      123166..138417
                  /note="assembly_fragment"
misc_feature      138518..151810
                  /note="assembly_fragment"
misc_feature      151911..156795
                  /note="assembly_fragment"
                  clone_end:T7
                  vector_side:right"
BASE COUNT      44911 a 31956 c 31627 g 45900 t 2401 others
ORIGIN
Alignment Scores:
Pred. No.:      0.223      Length:      156795
Score:          12.00      Matches:      12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      12.37%      Indels:      0
DB:              2      Gaps:      0
US-09-854-133-586 (1-97) x AC022290 (1-156795)
QY      23      ArgLYSLySGlUArgLYSLySArgGLUArgLYS 34
          |||||
Db 109212 AGAAGAAAGAAAGAAAGAAAGAGAGAGAAAGAAAG 109177
RESULT 30
AC024023/c      159875 bp      DNA      linear      PRI 09-JAN-2002
LOCUS      AC024023      Homo sapiens BAC clone RP11-83C7 from 4, complete sequence.
DEFINITION      AC024023
ACCESSION      AC024023
VERSION      AC024023.4      GI:18093064
KEYWORDS      HTG.
SOURCE      Homo sapiens.
              Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 159875)
AUTHORS      Sulston, J.E. and Waterston, R.
TITLE      Toward a complete human genome sequence
JOURNAL      Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE      99063792
PUBMED      9847074
REFERENCE      2 (bases 1 to 159875)
AUTHORS      Walligorski, J. and Cotton, M.
TITLE      The sequence of Homo sapiens BAC clone RP11-83C7
JOURNAL      Unpublished (2001)
REFERENCE      3 (bases 1 to 159875)
AUTHORS      Waterston, R.H.
TITLE      Direct Submission
JOURNAL      Submitted (20-FEB-2000) Genome Sequencing Center, Washington
              University School of Medicine, 4444 Forest Park Parkway, St. Louis,
              MO 63108, USA
              4 (bases 1 to 159875)
              Waterston, R.
              Direct Submission
              Submitted (09-JAN-2002) Department of Genetics, Washington
              University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
              On Jan 9, 2002 this sequence version replaced gi:15148356.
              ----- Genome Center
              Center: Washington University Genome Sequencing Center
              Center code: WUGSC
              Web site: http://genome.wustl.edu/gsc
              Contact: sapiens@wustl.edu
              ----- Summary Statistics
              -----
              Center project name: H_NH0083C07
              -----
NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
```

clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateo, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

#### VECTOR: PBACe3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is AC079088. Actual start of this clone is at base position 1 of RP11-83C7; actual end is at base position 159875 of RP11-83C7.

#### FEATURES

##### source

```
repeat_region      1..159875
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /chromosome="4"
                    /map="4"
                    /clone="RP11-83C7"
                    /clone_lib="RPCT-11"
                    17..1324
                    /rpt_family="MaLR"
                    1326..1497
                    /rpt_family="MaLR"
                    3301..3835
                    /rpt_family="ERV1"
                    3836..4139
                    /rpt_family="Alu"
                    3996..4027
                    /rpt_family="AT-rich"
                    4140..4414
                    /rpt_family="ERV1"
                    4413..4491
                    /rpt_family="ERV1"
                    4495..4578
                    /rpt_family="ERV1"
                    4871..8743
                    /rpt_family="L1"
                    6532..6557
                    /rpt_family="(CAA)n"
                    8498..8517
                    /rpt_family="(A)n"
                    8744..9292
                    /rpt_family="ERV1"
                    9293..9768
                    /rpt_family="L1"
                    9814..10894
                    /rpt_family="ERV1"
                    10895..11286
                    /rpt_family="L1"
                    11544..11956
                    /rpt_family="ERV1"
                    11957..12058
                    /rpt_family="ERV1"
```

Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: <http://www.nisc.nih.gov>  
Contact: nisc\_zooenhgri.nih.gov  
----- Project Information  
Center project name: czo  
Center clone name: 039N21  
----- Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 163467 bases at least Q40  
Consensus quality: 163520 bases at least Q30  
Consensus quality: 163523 bases at least Q20  
Insert size: 148000; agarose-gel



Insert size: 163780; sum-of-contigs  
Quality coverage: 11.14x in Q20 bases; agarose-fp  
Quality coverage: 10.07x in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 6103: contig of 6103 bp in length  
\* 6104 6203: gap of unknown length  
\* 6204 53231: contig of 47028 bp in length  
\* 53232 53331: gap of unknown length  
\* 53332 163980: contig of 110649 bp in length.

FEATURES  
Source Location/Qualifiers  
1. 163980  
/organism="Felis catus"  
/db\_xref="taxon:9685"  
/clone="RP86-39N21"  
/clone\_1b="RP86"

misc\_feature  
1. 6103  
/note="assembly\_fragment"

misc\_feature  
6204. 53231  
/note="assembly\_fragment"

misc\_feature  
clone\_end:SP6  
vector\_side:right  
53332. 163980  
/note="assembly\_fragment"

clone\_end:T7  
vector\_side:left

BASE COUNT 52242 a 28459 c 29919 g 53159 t 201 others  
ORIGIN

## Alignment Scores:

Pred. No.:	0.232	Length:	163980
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.37%	Indels:	0
DB:	2	Gaps:	0

US-09-854-133-586 (1-97) x AC120209 (1-163980)

QY 23 ArgLysLysGluArgLysLysArgGluArgLys 34  
|||||

Db 85437 AGAAGAAGAAAGAAAGAAAGAGAGAAAGAAAG 85472

RESULT 32  
AC023603/c 170513 bp DNA linear HTG 17-MAR-2000

LOCUS Homo sapiens chromosome 3p clone RP11-330014, WORKING DRAFT

DEFINITION SEQUENCE, 14 unordered pieces.

AC023603  
AC023603

AC023603.1 GI:6980169  
HTG; HTGS\_PHASE1; HTGS\_DRAFT.

KEYWORDS  
SOURCE  
ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
1 (bases 1 to 170513)  
Zhao, Y., Zhang, C., Lin, W., Dong, H., Wan, M., Xu, S., Gu, W., Tu, Y.,

Jia, J., Wu, C., Lu, G., Zhong, M., Zhou, Y., Ren, S., Fu, G., Chen, Z. and

Huang, M.

Direct Submission

Submitted (16-FEB-2000) Genomic Dept., Chinese National Human

Genome Center at Shanghai, 315 Guo Shou Jing Road, Shanghai,

Shanghai 201203, P. R. China

-----Genome Center Information-----  
Center: Chinese National Human Genome Center at Shanghai

Center Code: CHGC

Web site: <http://www.chgc.sh.cn>  
Email: [mhuang@chgc.sh.cn](mailto:mhuang@chgc.sh.cn) or [fugang@chgc.sh.cn](mailto:fugang@chgc.sh.cn)

-----End Genome Center Information-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 14 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 2230: contig of 2230 bp in length  
\* 2231 4248: contig of 2018 bp in length  
\* 4249 9512: contig of 5264 bp in length  
\* 9513 14048: contig of 4536 bp in length  
\* 14049 18094: contig of 4046 bp in length  
\* 18095 23536: contig of 5442 bp in length  
\* 23537 29239: contig of 5703 bp in length  
\* 29240 39461: contig of 10222 bp in length  
\* 39462 49383: contig of 9922 bp in length  
\* 49384 62126: contig of 12743 bp in length  
\* 62127 84920: contig of 22794 bp in length  
\* 84921 108776: contig of 23856 bp in length  
\* 108777 135846: contig of 27070 bp in length  
\* 135847 170513: contig of 34667 bp in length.

FEATURES  
Source Location/Qualifiers  
1. 170513  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="3p"  
/clone="RP11-330014"

BASE COUNT 53252 a 32291 c 31832 g 53096 t 42 others  
ORIGIN

## Alignment Scores:

Pred. No.:	0.24	Length:	170513
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.37%	Indels:	0
DB:	2	Gaps:	0

US-09-854-133-586 (1-97) x AC023603 (1-170513)

QY 23 ArgLysLysGluArgLysLysArgGluArgLys 34  
|||||

Db 119763 AGAAGAAGAAAGAAAGAAAGAGAGAAAGAAAG 119728

RESULT 33

AC120322/c 171506 bp DNA linear HTG 14-JUL-2002

LOCUS Rattus norvegicus clone CH230-313G4, \*\*\* SEQUENCING IN PROGRESS

DEFINITION \*\*\*, 64 unordered pieces.

AC120322  
AC120322.3 GI:21747387

AC120322.3 GI:21747387  
HTG; HTGS\_PHASE1.

KEYWORDS  
SOURCE  
ORGANISM

Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Muridae; Murinae;

REFERENCE  
AUTHORS  
Rattus.  
1 (bases 1 to 171506)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alshrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,  
Barbarta,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
DeLaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,R., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,  
Homsl,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlsoson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,  
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,  
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,  
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Plimus,E., Pu,L.L., Qules,M., Ren,Y.,  
Rives,M., Rojas,A., Rojuben,I., Rolfe,M., Ruiz,S., Savery,G.,  
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,  
Sodergren,E., Sonatke,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Taylor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.  
Direct Submission  
Unpublished  
2 (bases 1 to 171506)  
Worley,K.C.  
Direct Submission  
Submitted (06-MAY-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 171506)  
Worley,K.C.  
Direct Submission  
Submitted (14-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 14, 2002 this sequence version replaced gi:20799958.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GYZT  
Center clone name: CH230-313G4  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye: 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 110613 bases at least Q40  
Consensus quality: 116745 bases at least Q30  
Consensus quality: 121535 bases at least Q20

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 64 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1  
1085: contig of 1085 bp in length  
1185: gap of unknown length  
2230: contig of 1045 bp in length  
2330: gap of unknown length  
3772: contig of 1442 bp in length  
3872: gap of unknown length  
4940: contig of 1068 bp in length  
5040: gap of unknown length  
6080: contig of 1040 bp in length  
6180: gap of unknown length  
7584: contig of 1404 bp in length  
7685: gap of unknown length  
8945: contig of 1261 bp in length  
9045: gap of unknown length  
10666: contig of 1621 bp in length  
10766: gap of unknown length  
12113: contig of 1347 bp in length  
12213: gap of unknown length  
13776: contig of 1563 bp in length  
13876: gap of unknown length  
14948: contig of 1072 bp in length  
15048: gap of unknown length  
16293: contig of 1245 bp in length  
16393: gap of unknown length  
17761: contig of 1368 bp in length  
17861: gap of unknown length  
19757: contig of 1896 bp in length  
21148: gap of unknown length  
21248: contig of 1291 bp in length  
22701: contig of 1453 bp in length  
22801: gap of unknown length  
24254: contig of 1453 bp in length  
24354: gap of unknown length  
25664: contig of 1310 bp in length  
25764: gap of unknown length  
27300: contig of 1536 bp in length  
27400: gap of unknown length  
29535: contig of 2135 bp in length  
29635: gap of unknown length  
30916: contig of 1281 bp in length  
31016: gap of unknown length  
32517: contig of 1501 bp in length  
32617: gap of unknown length  
33883: contig of 1266 bp in length  
33983: gap of unknown length  
35888: contig of 1905 bp in length  
35988: gap of unknown length  
37153: contig of 1165 bp in length  
37253: gap of unknown length  
39294: contig of 2041 bp in length  
39394: gap of unknown length  
40996: contig of 1602 bp in length  
41096: gap of unknown length  
43689: contig of 2593 bp in length  
43789: gap of unknown length  
46865: contig of 3076 bp in length  
46965: gap of unknown length  
49295: contig of 2330 bp in length  
49395: gap of unknown length  
51847: contig of 2452 bp in length  
51947: gap of unknown length  
54027: contig of 2080 bp in length





misc_feature	35040. .39645
misc_feature	/note="assembly-fragment"
misc_feature	39746. .44851
misc_feature	/note="assembly-fragment"
misc_feature	44952. .50552
misc_feature	/note="assembly-fragment"
misc_feature	50653. .57894
misc_feature	/note="assembly-fragment"
misc_feature	57995. .65965
misc_feature	/note="assembly-fragment"
misc_feature	66066. .72594
misc_feature	/note="assembly-fragment"
misc_feature	72695. .80252
misc_feature	/note="assembly-fragment"
misc_feature	80353. .88942
misc_feature	/note="assembly-fragment"
misc_feature	89043. .96414
misc_feature	/note="assembly-fragment"
misc_feature	96515. .105678
misc_feature	/note="assembly-fragment"
misc_feature	105779. .115317
misc_feature	/note="assembly-fragment"
misc_feature	115418. .124960
misc_feature	/note="assembly-fragment"
misc_feature	125061. .135958
misc_feature	/note="assembly-fragment"
misc_feature	136059. .147620
misc_feature	/note="assembly-fragment"
misc_feature	147721. .160608
misc_feature	/note="assembly-fragment"
misc_feature	160709. .172571

BASE COUNT	vector_side:right"
53761	a 31812 c 31678 g 52614 t 2706 others
ORIGIN	

Alignment Scores:		
Pred. No.:	0.242	Length:
Score:	12.00	Matches:
Percent Similarity:	100.00%	Conservative:
Best Local Similarity:	100.00%	Mismatches:
Query Match:	12.37%	Indels:
DB:	2	0

US-09-854-133-586 (1-97) x AC027484 (1-172571)

Db 127661 AGAAGAGAAGAAAGAAAAAGAGAGAAAGCAG 127696

RESULT 35  
AC116823/c

DEFINITION	173700 bp	DNA	linear	HTG 12-JUL-2002
mus musculus clone RP24-365C10, 7 ordered pieces.			DRAFT SEQUENCE,	

ORGANISM	Mus musculus
SOURCE	house mouse.
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
ACCESSION	AC116823.3
GI	GI:21735055

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,  
Boukhgalter, B., Brown, A., Camarata, J., Campoliano, A., Chang, J.,  
Chazararo, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,  
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,





TITLE  
JOURNAL

## COMMENT

Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
Klein, J., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J.,  
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Margulis, N.,  
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,  
Meldrum, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,  
Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zahoun, J., Zimmer, A. and Zody, M.

Submitted (26-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 25, 2000 this sequence version replaced gi:7637290.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

Project Information

Center project name: L7466

Center clone name: 605\_A\_10

Summary Statistics

Sequencing vector: M13; M7815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 163013 bases at least Q40  
Consensus quality: 168440 bases at least Q30  
Insert size: 170000; agarose-ff  
Insert size: 171735; sum-of-ctigs  
Quality coverage: 4.6 in Q20 bases; agarose-ff  
Quality coverage: 4.5 in Q20 bases; sum-of-ctigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 21 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```

1 1198: contig of 1198 bp in length
* 1199 1298: gap of 100 bp
* 1299 2804: contig of 1506 bp in length
* 2805 2904: gap of 100 bp
* 2905 4901: contig of 1997 bp in length
* 4902 5001: gap of 100 bp
* 5002 8342: contig of 3341 bp in length
* 8343 8442: gap of 100 bp
* 8443 11290: contig of 2848 bp in length
* 11291 11390: gap of 100 bp
* 11391 15286: contig of 3896 bp in length
* 15287 15387: gap of 100 bp
* 15387 19353: contig of 3967 bp in length
* 19354 19453: gap of 100 bp
* 19454 24735: contig of 5282 bp in length
* 24736 24835: gap of 100 bp
* 24836 30442: contig of 5607 bp in length
* 30443 30542: gap of 100 bp
* 30543 38141: contig of 7599 bp in length
* 38142 38241: gap of 100 bp
* 38242 45256: contig of 7015 bp in length

```

```

* 45257 45356: gap of 100 bp
* 45357 52387: contig of 7031 bp in length
* 52388 52487: gap of 100 bp
* 52488 60034: contig of 7547 bp in length
* 60035 60134: gap of 100 bp
* 60135 68588: contig of 8454 bp in length
* 68589 68688: gap of 100 bp
* 68689 76018: contig of 7330 bp in length
* 76019 76118: gap of 100 bp
* 76119 86511: contig of 10393 bp in length
* 86512 86611: gap of 100 bp
* 86612 95770: contig of 9159 bp in length
* 95771 95870: gap of 100 bp
* 95871 107153: contig of 11283 bp in length
* 107154 107253: gap of 100 bp
* 107254 125495: contig of 18242 bp in length
* 125496 125595: gap of 100 bp
* 125596 144262: contig of 18667 bp in length
* 144263 144362: gap of 100 bp
* 144363 173735: contig of 29373 bp in length.

```

FEATURES  
source

```

misc_feature      /note="assembly-fragment"
                  1299..2804
                  /db_xref="taxon:9606"
                  /clone="RP11-605A10"
                  /clone_lib="RPCI-11 Human Male BAC"
                  1..1198
misc_feature      /note="assembly-fragment"
                  1299..2804
misc_feature      /note="assembly-fragment"
                  2905..4901
misc_feature      /note="assembly-fragment"
                  5002..8342
misc_feature      /note="assembly-fragment"
                  8443..11290
misc_feature      /note="assembly-fragment"
                  11391..15286
misc_feature      /note="assembly-fragment"
                  15387..19353
misc_feature      /note="assembly-fragment"
                  19454..24735
misc_feature      /note="assembly-fragment"
                  24836..30442
misc_feature      /note="assembly-fragment"
                  30543..38141
misc_feature      /note="assembly-fragment"
                  38242..45256
misc_feature      /note="assembly-fragment"
                  45357..52387
misc_feature      /note="assembly-fragment"
                  52488..60034
misc_feature      /note="assembly-fragment"
                  60135..68588
misc_feature      /note="assembly-fragment"
                  68689..76018
misc_feature      /note="assembly-fragment"
                  76119..86511
misc_feature      /note="assembly-fragment"
                  86612..95770
misc_feature      /note="assembly-fragment"
                  95871..107153
misc_feature      /note="assembly-fragment"
                  107254..125495
misc_feature      /note="assembly-fragment"
                  125596..144262
misc_feature      /note="assembly-fragment"
                  clone_end:SP6
                  vector_side:right"
                  144363..173735
                  /note="assembly-fragment"
BASE COUNT      52941 a 35819 c 34775 g 48197 t 2003 others
ORIGIN

```

Alignment Scores:  
Pred. No.: 0.244 Length: 173735  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
DB: 2 Gaps: 0

US-09-854-133-586 (1-97) x AC027063 (1-173735)

QY 8 HisAlaSerLeuGlyAspSerGluThrLeuSerGln 19  
|||||  
DB 159052 CATGCCAGCCTGGGGGACGTGAGACCCTGTCTCAA 159087

RESULT 37  
AC105990/c  
LOCUS  
DEFINITION Mus musculus clone RP24-175E7, WORKING DRAFT SEQUENCE, 4 ordered  
pieces.  
AC105990  
AC105990.4 GI:22381922  
HTG; HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_FULLTOP.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus  
house mouse  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
1 (bases 1 to 175398)  
Birren, B., Nusbaum, C. and Lander, E.  
Mus musculus, clone RP24-175E7  
Unpublished

2 (bases 1 to 175398)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,  
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,  
Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,  
Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,  
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K.,  
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,  
McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J.,  
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,  
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (11-JAN-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 175398)  
Birren, B., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S.,  
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,  
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,  
Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,  
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,  
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,  
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,  
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,  
McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,  
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,  
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P.,  
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,  
Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,

Viell, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,  
Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Aug 21, 2002 this sequence version replaced gi:20149427.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: 175\_E\_7  
Center clone name: L20237  
----- Summary Statistics  
Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 174705 bases at least Q40  
Consensus quality: 174895 bases at least Q30  
Consensus quality: 174990 bases at least Q20  
Insert size: 170000; agarose-fp  
Insert size: 175098; sum-of-contigs  
Quality coverage: 12.5 in Q20 bases; agarose-fp  
Quality coverage: 12.1 in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 4 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 17607: contig of 17607 bp in length  
\* 17608 17707: gap of 100 bp  
\* 17708 42885: contig of 25178 bp in length  
\* 42886 42985: gap of 100 bp  
\* 42986 66633: contig of 23648 bp in length  
\* 66634 66733: gap of 100 bp  
\* 66734 175398: contig of 108665 bp in length.  
-----  
Location/Qualifiers  
1. 175398  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP24-175E7"  
/clone\_lib="RPCI-24 Male Mouse BAC"  
1. 17607  
/note="assembly\_fragment"  
17708. . 42885  
/note="assembly\_fragment"  
42986. . 66633  
/note="assembly\_fragment"  
66734. . 175398  
/note="assembly\_fragment"

FEATURES  
source  
misc\_feature  
misc\_feature  
misc\_feature  
misc\_feature  
BASE COUNT 54842 a 32492 c 32477 g 55287 t 300 others  
ORIGIN

Alignment Scores:  
Pred. No.: 0.246 Length: 175398  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
DB: 2 Gaps: 0

US-09-854-133-586 (1-97) x AC105990 (1-175398)  
QY 23 ArgLysLysGluArgLysLysArgGluArgLys 34



Db 130324 AGAAAGAAAGAAAGAAAGAGAGAGAGAGAG 130289  
RESULT 38  
LOCUS AC009989/c 175690 bp DNA linear PRI 03-APR-2002  
DEFINITION Homo sapiens chromosome 10 clone RP11-7P17, complete sequence.  
ACCESSION AC009989  
VERSION AC009989.12 GI:19909351  
KEYWORDS HTG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS 1 (bases 1 to 175690)  
TITLE Smith, D.R.  
JOURNAL Genome Therapeutics Corporation Sequencing Center: Human Genome  
REFERENCE Unpublished  
AUTHORS 2 (bases 1 to 175690)  
TITLE Smith, D.R.  
JOURNAL Direct Submission  
REFERENCE Submitted (09-SEP-1999) Genome Therapeutics Corporation, 100 Beaver  
AUTHORS Street, Waltham, MA 02453, USA  
TITLE 3 (bases 1 to 175690)  
JOURNAL Smith, D.R.  
REFERENCE Direct Submission  
AUTHORS Submitted (03-APR-2002) Genome Therapeutics Corporation, 100 Beaver  
TITLE Street, Waltham, MA 02453, USA  
JOURNAL On Apr 3, 2002 this sequence version replaced gi:14670081.  
COMMENT  
FEATURES  
source  
1. 175690  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="10"  
/clone="RP11-7P17"  
/clone\_lib="RPCI-11"  
BASE COUNT 46780 a 36952 c 40536 g 51422 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.246 Length: 175690  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
DB: 9 Gaps: 0  
US-09-854-133-586 (1-97) x AC009989 (1-175690)  
QY 23 ArgLysLysGluArgLysLysLysArgGluArgLys 34  
Db 126623 AGAAAGAAAGAAAGAAAGAAAGAGAGAGAGAG 126588  
RESULT 39  
LOCUS AC011468/c 177444 bp DNA linear PRI 14-JUL-2002  
DEFINITION Homo sapiens chromosome 19 clone CTC-471J1, complete sequence.  
ACCESSION AC011468  
VERSION AC011468.8 GI:21747440  
KEYWORDS HTG.  
SOURCE human.  
ORGANISM Homo sapiens.  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS 1 (bases 1 to 177444)  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
JOURNAL DOE Joint Genome Institute and Stanford Human Genome Center.  
REFERENCE Unpublished  
AUTHORS 2 (bases 1 to 177444)  
TITLE DOE Joint Genome Institute.  
JOURNAL Direct Submission  
REFERENCE Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint

REFERENCE Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
AUTHORS 3 (bases 1 to 177444)  
TITLE DOE Joint Genome Institute and Stanford Human Genome Center.  
JOURNAL Direct Submission  
REFERENCE Submitted (23-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell  
AUTHORS Drive, Walnut Creek, CA 94598, USA  
TITLE 4 (bases 1 to 177444)  
JOURNAL DOE Joint Genome Institute and Stanford Human Genome Center.  
REFERENCE Direct Submission  
AUTHORS Submitted (14-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell  
TITLE Drive, Walnut Creek, CA 94598, USA  
JOURNAL On Jul 14, 2002 this sequence version replaced gi:15281205.  
COMMENT Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 99.9% of Sequence;  
Estimated Total Number of Errors is 0.1.  
FEATURES  
source  
1. 177444  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="CTC-471J1"  
BASE COUNT 55838 a 37182 c 36228 g 48196 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.248 Length: 177444  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
DB: 9 Gaps: 0  
US-09-854-133-586 (1-97) x AC011468 (1-177444)  
QY 8 HisAlaSerLeuGluYAspSerGluThrLeuSerGln 19  
Db 52413 CATGCCAGCCTGGGTGACAGTGAACCCCTCTCTCAA 52378  
RESULT 40  
LOCUS AC006042/c 178361 bp DNA linear PRI 21-DEC-1999  
DEFINITION Homo sapiens BAC clone RP11-505D17 from 7p22-p21, complete  
ACCESSION AC006042  
VERSION AC006042.2 GI:4508120  
KEYWORDS HTG.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
AUTHORS 1 (bases 1 to 178361)  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
JOURNAL Sulston, J.E. and Waterston, R.  
REFERENCE Toward a complete human genome sequence  
AUTHORS Genome Res. 8 (11), 1097-1108 (1998)  
TITLE 99063792  
JOURNAL PUBMED  
REFERENCE 9847074  
AUTHORS 2 (bases 1 to 178361)  
TITLE Hou, S., Wohlman, P. and Le, T.  
JOURNAL The sequence of Homo sapiens BAC clone RP11-505D17  
REFERENCE Unpublished  
AUTHORS 3 (bases 1 to 178361)  
TITLE Waterston, R.H.  
JOURNAL Direct Submission  
REFERENCE Submitted (22-NOV-1998) Genome Sequencing Center, Washington  
AUTHORS University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
TITLE MO 63108, USA  
JOURNAL 4 (bases 1 to 178361)  
REFERENCE Waterston, R.H.  
JOURNAL Direct Submission  
REFERENCE Submitted (24-MAR-1999) Genome Sequencing Center, Washington



REFERENCE  
 AUTHORS Waterston, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-AUG-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 6 (bases 1 to 178361)  
 WATERSTON, R.  
 REFERENCE  
 AUTHORS Direct Submission  
 TITLE Submitted (21-DEC-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 JOURNAL On Mar 24, 1999 this sequence version replaced g1:3980487.  
 COMMENT  
 ----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu/gsc>  
 Contact: [saplens@wustl.edu](mailto:saplens@wustl.edu)  
 ----- Summary Statistics  
 Center project name: H\_NH0505D17

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send [mailto:egreen@nhgri.nih.gov](mailto:mailto:egreen@nhgri.nih.gov), or see <http://genome.wustl.edu/gsc>

#### SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

#### VECTOR: pBACE3.6

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-438H20; the clone sequenced to the right is RP11-560C1. The actual start of this clone is at base position 1 of RP11-505D17; actual end is at base position 178361 of RP11-505D17.

#### Location/Qualifiers

1. 178361  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="7"  
 /map="7p22-p21"  
 /clone="RP11-505D17"  
 /clone\_lib="RPCI-11"  
 18. 267  
 /rpt\_family="L2"  
 334. 762  
 /rpt\_family="L2"  
 775. 802

#### FEATURES

##### source

repeat\_region  
 repeat\_region  
 repeat\_region

repeat\_region /rpt\_family="(CA)n"  
 817. 1043  
 /rpt\_family="L1"  
 misc\_feature 838. 1267  
 repeat\_region /note="match to EST AI025912 (NID:g3241525) ov95e10.s1"  
 1301. 1515  
 /rpt\_family="Alu"  
 2041. 2511  
 /note="match to EST AA759276 (NID:g2807139) ah90e10.s1"  
 5336. 5468  
 /rpt\_family="MIR"  
 repeat\_region 6287. 6449  
 /rpt\_family="L1"  
 repeat\_region 7747. 8030  
 /rpt\_family="Alu"  
 8032. 8058  
 /rpt\_family="(TAAA)n"  
 misc\_feature 8237. 8585  
 /note="match to EST AI223166 (NID:g3805369) q953d05.x1"  
 8241. 8581  
 /note="match to EST AA770501 (NID:g2821739) ah74b12.s1"  
 8248. 8649  
 /note="match to EST AA926996 (NID:g3075893) om26d02.s1"  
 9367. 9603  
 /rpt\_family="MALR"  
 repeat\_region 9604. 9896  
 /rpt\_family="Alu"  
 repeat\_region 9897. 10077  
 /rpt\_family="MALR"  
 repeat\_region 10680. 10981  
 /rpt\_family="Alu"  
 repeat\_region 10998. 11037  
 /rpt\_family="(CA)n"  
 11186. 11220  
 /rpt\_family="AT-rich"  
 repeat\_region 11587. 11713  
 /rpt\_family="MIR"  
 repeat\_region 11834. 12132  
 /rpt\_family="Alu"  
 repeat\_region 13508. 13607  
 /rpt\_family="(GAAA)n"  
 repeat\_region 13611. 13902  
 /rpt\_family="Alu"  
 repeat\_region 14757. 14932  
 /rpt\_family="MIR"  
 repeat\_region 16538. 16686  
 /rpt\_family="MER1-type"  
 repeat\_region 16819. 17280  
 /rpt\_family="MALR"  
 repeat\_region 17727. 18009  
 /rpt\_family="Alu"  
 repeat\_region 18104. 18313  
 /rpt\_family="L2"  
 repeat\_region 18487. 18974  
 /rpt\_family="L2"  
 repeat\_region 18975. 19261  
 /rpt\_family="Alu"  
 repeat\_region 19262. 20642  
 /rpt\_family="L2"  
 repeat\_region 21084. 21133  
 /rpt\_family="A-rich"  
 21186. 21766  
 /rpt\_family="L2"  
 repeat\_region 23013. 23147  
 /rpt\_family="Alu"  
 repeat\_region 23416. 23768  
 /rpt\_family="L2"  
 repeat\_region 25129. 25424  
 /rpt\_family="Alu"  
 repeat\_region 25553. 25603  
 /rpt\_family="AT-rich"  
 repeat\_region 25755. 25989  
 /rpt\_family="MIR"



ORIGIN

Alignment Scores:

Pred. No.:	0.25	Length:	178733
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.37%	Indels:	0
DB:	2	Gaps:	0

US-09-854-133-586 (1-97) x AC107305 (1-178733)

QY 8 HisAlaSerLeuGlyAspSerGluThrLeuSerGln 19  
|||||  
Db 48057 CATGCCAGCCTGGGGACAGTGAACCCCTGTCTCAA 48092

RESULT 42

AC121904/c 179343 bp DNA linear HTG 21-MAY-2002  
LOCUS Mus musculus chromosome UNK clone RP24-17507, WORKING DRAFT  
DEFINITION SEQUENCE, 7 unordered pieces.  
AC121904  
AC121904.1 GI:21040026  
HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
KEYWORDS Mus musculus.  
SOURCE Mus musculus  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 179343)  
McPherson, J.D. and Waterston, R.H.  
The sequence of Mus musculus clone  
Unpublished  
2 (bases 1 to 179343)  
McPherson, J.D. and Waterston, R.H.  
Direct Submission  
Submitted (21-MAY-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

COMMENT

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
Contact: submissions@wustl.wustl.edu  
----- Project Information -----  
Center project name: M\_BB0175007  
-----  
----- Summary Statistics -----  
Sequencing vector: M13; 0%  
Chemistry: Dye-primed ET; 0% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 175055 bases at least Q40  
Consensus quality: 175869 bases at least Q30  
Consensus quality: 176611 bases at least Q20  
Insert size: 157000; agarose-fp  
Insert size: 179794; sum-of-contigs  
Quality coverage: 0.00 in Q20 bases; agarose-fp  
Quality coverage: 13.00 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 7 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 4236: contig of 4236 bp in length  
\* 4237 4336: gap of unknown length  
\* 4337 10576: contig of 6240 bp in length  
\* 10577 10676: gap of unknown length  
\* 10677 31239: contig of 20563 bp in length

FEATURES

source

1.179343  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="UNK"  
/clone="RP24-17507"  
1.4236  
/note="assembly\_name:Contig28"  
4337.10576  
/note="assembly\_name:Contig29"  
10677.31239  
/note="assembly\_name:Contig30"  
31340.49523  
/note="assembly\_name:Contig31"  
49624.82058  
/note="assembly\_name:Contig32"  
82159.114218  
/note="assembly\_name:Contig33"  
114319.179343  
/note="assembly\_name:Contig34"

BASE COUNT 55906 a 34154 c 33514 g 55156 t 613 others

ORIGIN

Alignment Scores:

Pred. No.:	0.25	Length:	179343
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.37%	Indels:	0
DB:	2	Gaps:	0

US-09-854-133-586 (1-97) x AC121904 (1-179343)

QY 23 ArgLysLysGluArgLysLysLysArgGluArgLys 34  
|||||  
Db 4159 AGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4124

RESULT 43

AC108976/c 180336 bp DNA linear HTG 13-JUL-2002  
LOCUS Rattus norvegicus clone CH230-28717, \*\*\* SEQUENCING IN PROGRESS  
DEFINITION \*\*\* 62 unordered pieces.  
AC108976  
AC108976.3 GI:21737546  
HTG; HTGS\_PHASE1.  
KEYWORDS Norway rat.  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
1 (bases 1 to 180336)  
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Albrooks, S.L., Amaral, H.C., Are, J.R., Ayele, M., Banks, T., Barbara, J., Benton, J., Bimaye, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,

Correll, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B.,  
Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,  
Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
Massey, E., Mawhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,  
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
Nguyen, N., Nickerson, E., Nwokweto, S., Oguh, M., Okwuonu, G.,  
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savary, G.,  
Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,  
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
Weinstock, G., and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 180336)  
Worley, K.C.

Direct Submission  
Submitted (03-FEB-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 180336)  
Worley, K.C.

Direct Submission  
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 12, 2002 this sequence version replaced gi:18846553.

## COMMENT

Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
Project Information  
Center project name: GPIH  
Center clone name: CH230-28717  
Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 133122 bases at least Q40  
Consensus quality: 139518 bases at least Q30  
Consensus quality: 144903 bases at least Q20

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 62 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1  
\* 1179 1178: contig of 1178 bp in length  
\* 1279 1278: gap of unknown length  
\* 2490 2490: contig of 1212 bp in length  
\* 2491 2590: gap of unknown length  
\* 2591 3722: contig of 1132 bp in length  
\* 3723 3822: gap of unknown length

\* 3823 5011: contig of 1189 bp in length  
\* 5012 5111: gap of unknown length  
\* 5112 6446: contig of 1335 bp in length  
\* 6447 6546: gap of unknown length  
\* 6547 7769: contig of 1223 bp in length  
\* 7770 7869: gap of unknown length  
\* 7870 8905: contig of 1036 bp in length  
\* 8906 9005: gap of unknown length  
\* 9006 10193: contig of 1188 bp in length  
\* 10194 10293: gap of unknown length  
\* 10294 11596: contig of 1303 bp in length  
\* 11597 11696: gap of unknown length  
\* 11697 13543: contig of 1847 bp in length  
\* 13544 13643: gap of unknown length  
\* 13644 15725: contig of 2082 bp in length  
\* 15726 15825: gap of unknown length  
\* 15826 17704: contig of 1879 bp in length  
\* 15826 17804: gap of unknown length  
\* 17705 17805: contig of 1833 bp in length  
\* 17805 19637: gap of unknown length  
\* 19638 19737: gap of unknown length  
\* 19738 21484: contig of 1747 bp in length  
\* 21485 21584: gap of unknown length  
\* 21585 23839: contig of 2255 bp in length  
\* 23840 23939: gap of unknown length  
\* 23940 24969: contig of 1030 bp in length  
\* 24970 25070: gap of unknown length  
\* 25070 26477: contig of 1407 bp in length  
\* 26477 26577: gap of unknown length  
\* 26577 29205: contig of 2628 bp in length  
\* 29205 29305: gap of unknown length  
\* 29305 30860: contig of 1555 bp in length  
\* 30860 30959: gap of unknown length  
\* 30960 32587: contig of 1627 bp in length  
\* 32587 32687: gap of unknown length  
\* 32687 34481: contig of 1794 bp in length  
\* 34481 34580: gap of unknown length  
\* 34581 35936: contig of 1356 bp in length  
\* 35937 36037: gap of unknown length  
\* 36037 37576: contig of 1540 bp in length  
\* 37577 37676: gap of unknown length  
\* 37677 38964: contig of 1288 bp in length  
\* 38965 39065: gap of unknown length  
\* 39065 41372: contig of 2307 bp in length  
\* 41372 41472: gap of unknown length  
\* 41472 43482: contig of 2010 bp in length  
\* 43482 43581: gap of unknown length  
\* 43582 45439: contig of 1858 bp in length  
\* 45439 45539: gap of unknown length  
\* 45540 48069: contig of 2530 bp in length  
\* 48069 48169: gap of unknown length  
\* 48170 49876: contig of 1707 bp in length  
\* 49877 49976: gap of unknown length  
\* 49977 52658: contig of 2682 bp in length  
\* 52659 52758: gap of unknown length  
\* 52759 54635: contig of 1877 bp in length  
\* 54636 54735: gap of unknown length  
\* 54736 57622: contig of 2887 bp in length  
\* 57623 57722: gap of unknown length  
\* 57723 59661: contig of 1939 bp in length  
\* 59662 59761: gap of unknown length  
\* 59762 61953: contig of 2192 bp in length  
\* 61954 62053: gap of unknown length  
\* 62054 64139: contig of 2086 bp in length  
\* 64140 64239: gap of unknown length  
\* 64240 67756: contig of 3517 bp in length  
\* 67757 67856: gap of unknown length  
\* 67857 70906: contig of 3050 bp in length  
\* 70907 71006: gap of unknown length  
\* 71007 73236: contig of 2230 bp in length  
\* 73237 73336: gap of unknown length  
\* 73337 75973: contig of 2637 bp in length  
\* 75974 76073: gap of unknown length  
\* 76074 78833: contig of 2760 bp in length



```

* 78834 78933: gap of unknown length
* 78934 81146: contig of 2213 bp in length
* 81147 81246: gap of unknown length
* 81247 84514: contig of 3268 bp in length
* 84515 84614: gap of unknown length
* 84615 88398: contig of 3784 bp in length
* 88399 88498: gap of unknown length
* 88499 91447: contig of 2949 bp in length
* 91448 91548: gap of unknown length
* 91548 95023: contig of 3475 bp in length
* 95023 95122: gap of unknown length
* 95123 99168: contig of 4046 bp in length
* 99169 99269: gap of unknown length
* 99269 102681: contig of 3413 bp in length
* 102682 102781: gap of unknown length
* 102782 106572: contig of 3791 bp in length
* 106573 106673: gap of unknown length
* 106673 112143: contig of 5471 bp in length
* 112144 112243: gap of unknown length
* 112244 116064: contig of 3821 bp in length
* 116065 116164: gap of unknown length
* 116165 120276: contig of 4112 bp in length
* 120277 120376: gap of unknown length
* 120377 123959: contig of 3583 bp in length
* 123960 124059: gap of unknown length
* 124060 128058: contig of 3999 bp in length
* 128059 128158: gap of unknown length
* 128159 132524: contig of 4366 bp in length

```

## Alignment Scores:

```

Pred. No.: 0.252 Length: 180336
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.37% Indels: 0
DB: 2 Gaps: 0

```

US-09-854-133-586 (1-97) x AC108976 (1-180336)

QY 23 ArgLysLysGluArgLysLysArgGluArgLys 34

Db 63702 AGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAG 63667

## RESULT 44

AC026367/c

LOCUS

DEFINITION Homo sapiens 12 BAC RP11-131L12 (Roswell Park Cancer Institute

AC026367 Human BAC Library) complete sequence.

AC026367 AC026367.14 GI:14669931

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 181445)  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 Muszay,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbara,J.,  
 Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,  
 Bowle,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,  
 Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,  
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,  
 Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,  
 Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,  
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,  
 Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,  
 Dugan-Rocha,S., Durbin,K.J., Eamhart,C., Edgar,D., Edwards,C.C.,  
 Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D.,  
 Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,  
 Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W.,  
 Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K.,  
 Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O.,  
 Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homs,F.,

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 JOURNAL

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL

## COMMENT

Howard,S., Huber,J., Eulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C.,  
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,  
 Louised,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
 Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R.,  
 Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P.,  
 Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G.,  
 Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M.,  
 Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N.,  
 Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,  
 Ogun,M., Okunonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
 Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,  
 Qules,M., Ren,Y., Rives,M., Rojas,A., Rojudoan,I., Rolfe,M.,  
 Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C.,  
 Shooshtari,N., Sisson,I., Sodergren,E., Sonake,T., Sparks,A.,  
 Stanley,H., Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A.,  
 Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B.,  
 Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D.,  
 Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R.,  
 Washington,C., Watlington,S., Williams,G., Williamson,A.,  
 Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J.,  
 Zorrilla,S., Kucherlapati,R. and Gibbs,R.

Direct Submission  
 Unpublished  
 2 (bases 1 to 181445)  
 Worley,K.C.

Direct Submission  
 Submitted (22-MAR-2000) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 181445)  
 Worley,K.C.

Direct Submission  
 Submitted (11-JUL-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA

On Jul 11, 2001 this sequence version replaced gi:13493004.  
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
[gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the  
 entire insert of this clone. Overlapping regions of clones are only  
 sequenced and submitted once, so the sequence for the remainder of  
 the insert may be found in the record for the adjacent clones.  
 Overlapping clones are noted at the beginning and end of the  
 Features listing.

## ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches  
 of a local database that includes entries from dbSTS, GDB, and  
 local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,  
 unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST  
 (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the  
 EST and cDNA sequences. Genes demonstrate at least two exons  
 flanked by consensus splice sites that maintained sequence  
 continuity across the splice junctions. Sequences that are not  
 identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum  
 standard of double strand coverage with a minimum of 2 clones and 2  
 reads with no ambiguities or 2 chemistries with a minimum of 2  
 clones and 3 reads with no ambiguities. If the sequence quality for  
 a region does not meet this standard, it will be indicated in the  
 annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality  
 standards - estimated error rate less than 1 per 10,000 bases.  
 Reports of lowest quality individual bases and measures of base  
 quality are listed below. Description of the metrics can be found

at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

## QUALSTAT-REPORT-----

## ----- Summary Statistics -----

Contig length: 181445  
 Phrap values in estimate: 180115  
 Average error rate (BCM-Phrap estimate): 0.000170582  
 Fraction of Phrap values less than 40 : 0.0184993  
 Number of consensus changing edits: 46  
 Number of N's in consensus : 0

Position	Consensus changing edits	Original+Context	Edited+Context
1715	aggctggtct(n)gacctcctga	aggctggtct(c)gacctcctga	
1743	tgatccacc(n)ccttgccctc	tgatccacc(a)ccttgccctc	
1806	ggagtgtatga(n)tattttaaa	ggagtgtatga(t)tattttaaa	
1843	gtacatgccc(n)atgtagaana	gtacatgccc(a)atgtagaana	
1852	cnatgtagaa(n)acttgggaaa	cnatgtagaa(a)acttgggaaa	
1866	tggaataac(n)cattatgaag	tggaataac(a)cattatgaag	
1909	tggtcctaaa(n)ngtaataccc	tggtcctaaa(c)ngtaataccc	
1910	ggtcctaaa(n)ngtaataccc	ggtcctaaa(c)ngtaataccc	
1911	ggtcctaaa(n)ngtaataccc	ggtcctaaa(c)ngtaataccc	
2049	gtgcctgtaa(n)cccagctact	gtgcctgtaa(t)cccagctact	
2067	actcaggagg(a)tgaggcagga	actcaggagg(c)tgaggcagga	
2114	gttgcaagtga(n)ccagatnng	gttgcaagtga(g)ccagatnng	
2122	gancaggat(n)gncaccatgc	gancaggat(c)gncaccatgc	
2125	ccgagatnng(n)ccactgcnct	ccgagatnng(a)ccactgcnct	
2133	gncaccatgc(n)ctccagcctg	gncaccatgc(a)ctccagcctg	
2163	gtgagactcc(n)tctcaaaaaa	gtgagactcc(a)tctcaaaaaa	
4307	atgtgtcct(n)taagatttc	atgtgtcct(c)taagatttc	
36061	gagggcggaa(n)ntncaagtga	gagggcggaa(g)ntncaagtga	
36062	gagggcggaa(n)ntncaagtga	gagggcggaa(g)ntncaagtga	
36064	ggcggaaant(n)cagtgaagtg	ggcggaaant(g)cagtgaagtg	
51326	gtgggtgat(n)gcttgaagntc	gtgggtgat(t)gcttgaagntc	
51334	atngctlgag(n)tcaggagatc	atngctlgag(c)tcaggagatc	
51584	ctgggagaa(n)ttttttttt	ctgggagaa(t)ttttttttt	
52661	aggatggtct(n)gatttctaga	aggatggtct(c)gatttctaga	
52675	ttctagacct(n)gtgactctgc	ttctagacct(c)gtgactctgc	
57080	cacaatctca(n)ctaactgcaa	cacaatctca(g)ctaactgcaa	
64017	aataaaaaa(n)ggagacatga	aataaaaaa(a)ggagacatga	
86177	tttttttag(n)ggcggagctc	tttttttag(t)ggcggagctc	
92345	cattttatt(n)aacctcattc	cattttatt(a)aacctcattc	
97328	aaaaaaaaa(n)gaaagaagaag	aaaaaaaaa(a)gaaagaagaag	
102654	gctaagtgt(n)tatttttagt	gctaagtgt(g)tatttttagt	
102701	aggctggtct(n)gnattcctga	aggctggtct(c)gnattcctga	
102703	gctggtctng(n)atcctgact	gctggtctng(a)atcctgact	
102718	ctgacttcag(n)ngattcaacc	ctgacttcag(g)ngattcaacc	
102719	tgacttcagn(n)gattcaaccg	tgacttcag(t)gattcaaccg	
102784	ggccggcctg(n)natTTTTTT	ggccggcctg(g)natTTTTTT	
102785	cccgccctgn(n)natTTTTTT	cccgccctgn(t)natTTTTTT	
102786	ccggcctgm(n)atTTTTTTT	ccggcctgm(t)atTTTTTTT	
102820	ccgggagcgg(n)ggtcaacgcc	ccgggagcgg(t)ggtcaacgcc	
113291	agatggggt(n)nnanacngng	agatggggt(c)nnanacngng	
113292	gatggggtt(n)annacngngg	gatggggtt(c)annacngngg	
113294	tggggttma(n)nacngnggca	tggggttma(c)nacngnggca	
113295	gggggttma(n)acngnggca	gggggttma(c)acngnggca	
113299	ttmanacng(n)ggccagcgt	ttmanacng(t)ggccagcgt	
113300	ttmanacng(n)ggccagcgt	ttmanacng(t)ggccagcgt	
155971	ttgggacatt(n)tcttagtctc	ttgggacatt(t)tcttagtctc	

----- Distribution of Quality < 40 Bases -----

#	bases
10001	*
9001	*
8001	*
7001	*
6001	*
5001	*

	5	10	15	20	25	30	35	40
4001							*	*
3001							*	*
2001			*	*	*	*	*	*
1001		*	*	*	*	*	*	*
01	*	*	*	*	*	*	*	*

Phrap Value Range

Version: 1.01 gxf.  
 Location/Qualifiers  
 1.181445

FEATURES  
 Source  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="12"  
 /clone="RP11-131L12"  
 complement(469..782)  
 /rpt\_family="AluSp"  
 complement(1499..1795)  
 /rpt\_family="AluSx"  
 1887..2197  
 /rpt\_family="AluSg"  
 3011..3319  
 /rpt\_family="AluSp"

## Alignment Scores:

Pred. No.: 0.253  
 Score: 12.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 12.37%  
 DB: 9  
 Gaps: 0

US-09-854-133-586 (1-97) x AC026367 (1-181445)

QY 8 HisAlaSerLeuGlyAspSerGluThrLeuSerGln 19

Db 141234 CATGCCAGCTGGGGGACAGTGAGACCCCTGTCTCAA 141199

RESULT 45  
 AC118757/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

TITLE

JOURNAL

REFERENCE

AUTHORS

AC118757 183011 bp DNA linear HTG 23-AUG-2002  
 Homo sapiens chromosome 18 clone RP11-905K4 map 18, \*\*\* SEQUENCING  
 IN PROGRESS \*\*\*, 2 ordered pieces.  
 AC118757  
 AC118757.8 GI:22450658  
 HTG; HTGS\_PHASE2; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
 human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 183011)  
 Birren,B., Nussbaum,C. and Lander,E.  
 Homo sapiens chromosome 18, clone RP11-905K4  
 Unpublished  
 2 (bases 1 to 183011)  
 Birren,B., Linton,L., Nussbaum,C., Lander,E., Ali,A., Allen,N.,  
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,  
 Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,  
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,  
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
 Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,  
 Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,  
 Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
 McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Menus,L.,  
 Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,  
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
COMMENT  
Direct Submission  
Submitted (20-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 183011)  
Birren,B., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L26244  
Center clone name: 905\_K\_4

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submittor.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 116580: contig of 116580 bp in length  
\* 116581 116680: gap of 100 bp  
\* 116681 183011: contig of 66331 bp in length.  
Location/Qualifiers  
1.183011  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="18"  
/map="18"  
/clone="RP11-905K4"  
/clone\_lib="RPC1-11 Human Male BAC"

BASE COUNT 49244 a 37615 c 38929 g 56894 t 329 others  
ORIGIN

Alignment Scores:  
Pred. No.: 0.255 Length: 183011  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 12.37% Indels: 0  
DB: 2 Gaps: 0  
US-09-854-133-586 (1-97) x AC118757 (1-183011)  
QY 23 ArgLysLysGluArgLysLysLysArgGluArgLys 34  
|||||  
Db 7219 AGAAGAAGAAAGAAAGAAAGAAAGAAAGAAAG 7184

Search completed: May 11, 2003, 17:19:47  
Job time : 2680.93 secs





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 11, 2003, 16:14:30 ; Search time 207.735 Seconds  
(without alignments)  
1051.553 Million cell updates/sec

Title: US-09-854-133-586

Perfect score: 97  
Sequence: 1 EVEVSRDHASLGDESETLSQT.....LTGGCLPWATRSHLGRKCS 97

Scoring table:

OLIGO	
Xgapop 60.0 , Xgapext 60.0	
Ygapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2185239 seqs, 1125999159 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4368722

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 60 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09854133/runat\_05052003\_174132\_694/app\_query.fasta\_1.462  
-DB=N\_Geneseq\_101002 -QFMT=fastap -SUFFIX=oligo.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=60 -DOCALIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=45 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09854133\_ECGN\_1\_1\_275\_@runat\_05052003\_174132\_694 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAL -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002.\*

1:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24:	/SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	100.0	337	23 AAD23462	Human lung tumour-
2	96	99.0	5981	23 AAD23461	Human lung tumour-
3	93	95.9	2239	23 AAD23460	Human lung tumour-
4	37	38.1	575	24 ABL37718	Human colon tumour
5	37	38.1	596	24 ABK27651	Human gene express
6	37	38.1	800	20 AAZ16609	Human cysteine/Glu
7	37	38.1	1958	22 ABA09201	Human EST-derived
8	37	38.1	123	22 AAK68639	Human muscleoskele
9	36	37.1	29607	22 AAL37261	Human cardiovascular
10	12	11.4	32190	22 AAS36709	Human immune/haema
11	11	11.3	136	22 AAK82863	Human immune/haema
12	10	10.3	136	22 AAK82864	Human secreted pro
13	10	10.3	207	21 AAC23449	Human secreted exp
14	10	10.3	291	21 AAA44648	Novel human diagno
15	10	10.3	291	22 AAS37013	Novel human diagno
16	10	10.3	291	22 AAS37176	Human gene express
17	10	10.3	299	22 AAK75253	Human gene express
18	10	10.3	300	20 AAZ14169	Human gene express
19	10	10.3	300	20 AAZ13802	Human gene express
20	10	10.3	300	20 AAZ13190	Human gene express
21	10	10.3	303	22 AAS37025	Novel human diagno
22	10	10.3	310	22 AAK75254	Human gene express
23	10	10.3	343	22 AA192003	Human immune/haema
24	10	10.3	351	24 ABL63133	Human polynucleoti
25	10	10.3	371	23 ABV19284	Breast cancer rela
26	10	10.3	376	23 ABV13690	Human prostate exp
27	10	10.3	394	22 ABV13690	Human prostate exp
28	10	10.3	401	23 ABV04521	Human prostate exp
29	10	10.3	401	23 ABV34803	Human prostate exp
30	10	10.3	401	23 ABV43654	Human ovarian canc
31	10	10.3	418	24 ABL84142	Human polynucleoti
32	10	10.3	430	22 AA192526	Human prostate exp
33	10	10.3	440	23 ABV49058	Human prostate exp
34	10	10.3	457	22 AAH03381	Human CDNA clone (
35	10	10.3	459	22 AA198951	Human excretory re
36	10	10.3	459	22 AA163301	Human kidney relat
37	10	10.3	486	23 ABV49991	Human prostate exp
38	10	10.3	490	23 ABV46023	Human prostate exp
39	10	10.3	491	23 ABV49206	Human prostate exp
40	10	10.3	497	22 AAS30313	Human encoding novel
41	10	10.3	497	22 AAL01994	Human reproductive
42	10	10.3	509	22 AAH09380	Human CDNA clone (
43	10	10.3	520	22 ABA59875	Human foetal liver
44	10	10.3	520	22 AAK08145	Human brain expres
45	10	10.3	520	22 AAK34023	Human bone marrow
46	10	10.3	523	22 AAK88938	Human digestive sy
47	10	10.3	525	22 AAH09825	Human CDNA clone (
48	10	10.3	527	22 AAK85814	Human CDNA clone (
49	10	10.3	550	22 AAH04912	Human CDNA clone (
50	10	10.3	550	22 AAH11147	Human prostate exp
51	10	10.3	586	23 ABV19433	Human immune/haema
52	10	10.3	642	22 AAK70257	Human inflammatory
53	10	10.3	700	22 AAH92254	Human inflammatory
54	10	10.3	700	22 AAH92255	Human inflammatory
55	10	10.3	700	22 AAH92301	Human inflammatory
56	10	10.3	700	22 AAH92302	Human inflammatory
57	10	10.3	771	22 AA195159	Human neuroblastom
58	10	10.3	868	22 AAK80200	Human immune/haema
59	10	10.3	868	22 AAK82572	Human immune/haema
60	10	10.3			

## ALIGNMENTS

RESULT 1  
AAD23462  
ID AAD23462 standard; cDNA; 337 BP.

XX	AAD23462;	
AC		
XX	26-FEB-2002	(first entry)
DT		
XX	Human lung tumour-specific	20E10 5' CDNA.
DE		
XX	Human; lung tumour protein; immunostimulant; cytostatic; gene therapy; antisense-therapy; vaccine; immune response; lung cancer; 20E10; ss.	
KW		
KW		
XX		
OS	Homo sapiens.	

WO200172295-A2

04-OCT-2001.

48-MAR-2001; 2001WO-US099991.

25-MAR-2000; 2000US-0538037.  
05-JUN-2000; 2000US-0588937.  
18-AUG-2000; 2000US-0640878.  
22-SEP-2000; 2000US-234517P.  
01-NOV-2000; 2000US-0704512.  
14-DEC-2000; 2000US-0738973.

(CORI-) CORIXA CORP.

WPI: 2001-639201/73.

diagnosis and treatment of disease e.g. lung cancer -  
 Claim 1; Page 334; 378pp; English.

CC The invention relates to isolated lung tumour-specific proteins and  
CC their corresponding cDNA molecules. Lung tumour-specific proteins and  
CC their antigen-presenting cells are useful for stimulating and/or  
CC expanding T cells specific for a tumour protein, and for inhibiting  
CC the development of cancer. The invention also relates to a composition  
CC useful for stimulating an immune response, and for treating cancer. The  
CC lung tumour specific oligonucleotide is useful in gene therapy and for  
CC diagnosis, detection and treatment of lung cancer. The present sequence  
CC is a cDNA encoding human lung tumour-specific protein.  
XX

reference 55; BF; 103 A; 60 C; 93 G; 81 T; 0 other;

Alignment Scores:  
Pred. No.:

freq. NO.:	1.22e-89	Length:	337
Score:	97.00	Matches:	97
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	23	Gaps:	0

US-09-854-133-586 (1-97) x AAD23462 (1-337)

QY 1 GluValGluValSerArgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThr 20  
 |||  
 Db 5 GAGCTGAAGTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAACA 64  
 |||

QY 21 G I L E u A r g l y s L y s G l u A r g L y s L y s A r g G l u A r g L y s P h e G l n A l a S n C y s G l y 40  
 |||||  
 65 G A A T T A A G C A A A A A G A A A G A A A G A G A G A G A A A T T C C A G G C C A A T T G T G C C 120  
 |||||

[illegible]

61 GluSerLeuLeuCysProProSerProLysGluValThrCysArgGluMetLeuThrGly 80  
|||||

D <sub>b</sub>	185	GAAAGCCTGTGTTGGTCCACCATCTCCAAGAAGGATTACCTGCAGGAATA GTTAACGGGA	244
OY	81	GLYCysleuproTrpalathrargserHisleuglyArgglyscySer	97
D <sub>b</sub>	245	GGCTGCCCTTCCTGGGCAACAAGAGCCACTGGGCAGAGAGAAGTGCCAGC	295

RESULT 2  
AAD23461

ID	standard; CDNA; 5981 BP.
XX	AAD23461

AC AAD23461;  
VY

26-FEB-2002 (first entry)

XX tumour-specific 14F10 full length cDNA.

KW Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;  
KW antisense-therapy; vaccine; immune response; lung cancer; 14F10; ss.  
XX  
OS Homo sapiens.  
XX

W0200172295-A2.

04-OCT-2001

28-MAR-2001; 2001WO-US09991.

05-MAR-2000; 2000US-0538037  
05-JUN-2000; 2000US-0588937  
18-AUG-2000; 2000US-0640878  
22-SEP-2000; 2000US-2734517P  
01-NOV-2000; 2000US-0304512  
14-DEC-2000; 2000US-0738973

(CORI-) CORIXA CORP.

1 Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indrias CY;  
I Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;  
R WPI; 2001-639201/73.

New human lung-specific polynucleotides and polypeptides for the diagnosis and treatment of disease e.g. lung cancer -  
Claim 1; Page 332-334; 378pp; English.

The invention relates to isolated lung tumour-specific proteins and their corresponding cDNA molecules. Lung tumour-specific proteins and their antigen-presenting cells are useful for stimulating and/or expanding T cells specific for a tumour protein, and for inhibiting the development of cancer. The invention also relates to a composition useful for stimulating an immune response, and for treating lung tumour specific oligonucleotide is useful in gene therapy and for diagnosis, detection and treatment of lung cancer. The present sequence is a cDNA encoding human lung tumour-specific protein.

Reference 2001 BF; 1838 A; 1029 C; 1098 G; 1996 T; 0 other;

**Alignment Scores:**

tree. NO.:	1.85e-87	Length:	5981
Score:	96.00	Matches:	96
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.97%	Indels:	0
DB:	23	Gap:	0

OS-09-854-133-586 (1-97) x AAD23461 (1-5981)

2 ValGluValSerArgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThrGln 31

DB 3 GTGAGCTGAGCAGAGATCATGCCAGCCTGGGTGACAGTGAGACTCTGTCTCAACGAA 62

QY 22 LeuArgLysLysGluArgLysLysLysArgGluArgLysPheGlnAlaAsnCySylIle 41

```
DB: 63 TTAAGGAAAAAGAAAGAAAAAGAGAGAGAAATTCAGGCAATTGTGCATA 122
QY: 42 AspHeileleIephetrIlePheTrIleuLeuPheSerHisTrIleGlnGlu 61
DB: 123 GATTTATCATATCTGTGATTTTGGATCTTTGTTTCTCATCTGATTCAGAA 182
QY: 62 SerLeuLeuCySProProSerProLySGluValThrCySArgGluMetLeuThrGly 81
DB: 183 AGCCTGTGTGTCCACCATCTCCAAAGAGGTTACCTGCAGGAAATGTTAAAGGAGGC 242
QY: 82 CysLeuProTrpAlaThrArgSerHisLeuGlyArgArgLysCysSer 97
DB: 243 TGCCTTCCCTGGGCAACAAGAGCCACCTGGGCAAGAGAAAGTGACAGC 290

RESULT 3
AAB23460
ID: AAD23460 standard; cDNA; 2239 BP.
XX:
AC: AAD23460;
DT: 26-FEB-2002 (first entry)
DE: Human lung tumour-specific 19A4 cDNA.
KW: Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
    antisense-therapy; vaccine; immune response; lung cancer; 19A4; ss.
OS: Homo sapiens.
XX:
PN: WO200172295-A2.
XX:
PD: 04-OCT-2001.
PF: 28-MAR-2001; 2001WO-US09991.
XX:
PR: 29-MAR-2000; 2000US-0538037.
PR: 05-JUN-2000; 2000US-0588937.
PR: 18-AUG-2000; 2000US-0640878.
PR: 22-SEP-2000; 2000US-234517P.
PR: 01-NOV-2000; 2000US-0704512.
PR: 14-DEC-2000; 2000US-0738973.
XX:
PA: (CORI-) CORIXA CORP.
XX:
PI: Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indirias CY;
    Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
XX:
DR: WPI; 2001-639201/73.
XX:
PT: New human lung-specific polynucleotides and polypeptides for the
    diagnosis and treatment of disease e.g. lung cancer -
XX:
PS: Claim 1; Page 332; 378pp; English.
XX:
CC: The invention relates to isolated lung tumour-specific proteins and
    their corresponding cDNA molecules. Lung tumour-specific proteins and
    their antigen-presenting cells are useful for stimulating and/or
    expanding T cells specific for a tumour protein, and for inhibiting
    the development of cancer. The invention also relates to a composition
    useful for stimulating an immune response, and for treating cancer. The
    lung tumour specific oligonucleotide is useful in gene therapy and for
    diagnosis, detection and treatment of lung cancer. The present sequence
    is a cDNA encoding human lung tumour-specific protein.
XX:
SQ: Sequence 2239 BP; 619 A; 444 C; 493 G; 683 T; 0 other;

Alignment Scores:
Pred. No.: 8.3e-85 Length: 2239
Score: 93.00 Matches: 93
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.88% Indels: 0
```

```
DB: 23 Gaps: 0
US-09-854-133-586 (1-97) x AAD23460 (1-2239)
QY: 1 GluValGluValSerArgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThr 20
DB: 2 GAGGTGAAGTGAGCAGAGATCATGCGACCCCTGGGTGACAGTGAAGACTGTCTCAACA 61
QY: 21 GluLeuArgLysLysGluArgLysLysLysArgGluArgLysPheGlnAlaAsnCysGly 40
DB: 62 GAAATTAAAGGAAAAAGAAAGAAAAAGAGAGAGAGAAATTCAGGCAATTGTGCGC 121
QY: 41 IleAspHeileleIephetrIlePheTrIleuLeuPheSerHisTrIleGln 60
DB: 122 ATAGATTTTATCATATCTGTGATTTTGGATTCCTTTGTTTCTCATCTGATTCAG 181
QY: 61 GluSerLeuLeuCySProProSerProLySGluValThrCySArgGluMetLeuThrGly 80
DB: 182 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGAGGTTACCTGCAGGAAATGTTAACGGGA 241
QY: 81 GlyCysLeuProTrpAlaThrArgSerHisLeuGlyArg 93
DB: 242 GGCTGCCCTTCCCTGGGCAACAAGAGCCACCTGGGCAAG 280

RESULT 4
ABL37718
ID: ABL37718 standard; cDNA; 575 BP.
XX:
AC: ABL37718;
DT: 08-APR-2002 (first entry)
DE: Human colon tumour antigen polynucleotide SEQ ID NO:1307.
KW: Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
    colon tumour metastatic antigen; diagnosis; gene; ss.
OS: Homo sapiens.
XX:
PN: WO200196388-A2.
XX:
PD: 20-DEC-2001.
PF: 08-JUN-2001; 2001WO-US18557.
XX:
PR: 09-JUN-2000; 2000US-210899P.
PR: 20-FEB-2001; 2001US-270216P.
XX:
PA: (CORI-) CORIXA CORP.
XX:
PI: Jiang Y, Harlocker SL, Secrist H;
XX:
DR: WPI; 2002-114514/15.
XX:
PT: Novel isolated colon tumor polynucleotide differentially expressed in
    colon tumor or colon metastatic tumor and polypeptides encoded by them,
    useful for inhibiting development of cancer in patient -
XX:
PS: Claim 1; SEQ ID 1307; 105pp; English.
XX:
CC: ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
    which were isolated from human colon tumour and colon metastatic tumour
    cDNA libraries. (I) have cytostatic activity and can be used in vaccine
    production. (I) can be used for stimulating and/or expanding T cells
    specific for a tumour protein on contact with the T cells. They are also
    useful for inhibiting the development of cancer in a patient. (I) can be
    used as probes or primers for nucleic acid hybridisation, for preparing
    mutant species primers, or primers for use in genetic constructions. (I)
    can be used in the diagnosis of a colon tumour.
XX:
SQ: Sequence 575 BP; 151 A; 122 C; 153 G; 141 T; 8 other;

Alignment Scores:
```

Pred. No.: 1.9e-28  
Score: 37.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 38.14%  
DB: 24  
Length: 575  
Matches: 37  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x ABL37718 (1-575)

QY 61 GluSerLeuLeuCySProProSerProLysGluValThrCysArgGluMetLeuThrGly 80  
DB 256 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGGAGTTACCTGCAGGAAATGTTACGGGA 315  
QY 81 GLYCysLeuProTTPalaThrArgSerHisLeuGlyArgArgLysCysSer 97  
DB 316 GGCTGCCCTTCCCTGGGCAACAAGAGCCACCTGGGCGAGAGAAAGTGACAGC 366

## RESULT 5

ABK27651/c  
ID ABK27651 standard; cDNA; 596 BP.  
XX  
AC ABK27651;  
XX  
DT 09-APR-2002 (first entry)  
XX  
DE Human colon cancer expressed sequence tag, Seq ID no 88.  
XX  
KW Human; colon cancer; T cell expansion; tumour; EST; gene; ss;  
XX  
OS Homo sapiens.  
XX  
PN WO200196390-A2.  
XX  
PD 20-DEC-2001.  
XX  
PF 08-JUN-2001; 2001WO-US18577.  
XX  
PR 09-JUN-2000; 2000US-210821P.  
PR 18-DEC-2000; 2000US-256571P.  
PR 10-MAY-2001; 2001US-290240P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Jiang Y, Hepler WT, Clapper JD, Wang A, Secrist H;  
XX  
DR WPI; 2002-139708/18.  
XX

Novel isolated polynucleotide encoding a polypeptide comprising a portion of colon tumour protein, useful for detection, diagnosis and therapy of human colon cancer

Claim 1; Page 174; 220pp; English.

The invention relates to an isolated polynucleotide (I) encoding a polypeptide (II) comprising at least a portion of a colon tumour protein. (I), (II) and antibody (III) to (II) are useful for determining the presence of a cancer in a patient. (I), (II) or antigen presenting cells expressing (I) is useful for stimulating and/or expanding T cells specific for a tumour protein, by contacting T cells with (I), (II) or antigen-presenting cells that express (I), under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells. (I), (II), or antigen presenting cells that express (II) are useful for treating colon cancer in a patient by incubating CD4+ and/or CD8+ T cells isolated from a patient with (I), (II) or antigen presenting cells that express (II), such that T cells proliferate, and administering to the patient an effective amount of the proliferated T cells, thus inhibiting the development of a cancer in the patient. (I) or (II) is useful in vaccines and pharmaceutical compositions for prevention and treatment of colon malignancies and for the diagnosis and monitoring of such cancers. (I), (II) or (III) is useful for detection, diagnosis and/or therapy of human colon cancer. (I) is useful as a probe or primer for nucleic acid hybridisation, and in the design and preparation of

CC ribozyme molecules for inhibiting expression of (II) in tumour cells.  
CC ABK27564-ABK27807 represent novel human colon cancer coding  
CC sequences and primers of the invention.  
XX  
SQ Sequence 596 BP; 160 A; 151 C; 126 G; 158 T; 1 other;

## Alignment Scores:

Pred. No.: 1.97e-28  
Score: 37.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 38.14%  
DB: 24  
Length: 596  
Matches: 37  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x ABK27651 (1-596)

QY 61 GluSerLeuLeuCySProProSerProLysGluValThrCysArgGluMetLeuThrGly 80  
DB 367 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGGAGTTACCTGCAGGAAATGTTACGGGA 308  
QY 81 GLYCysLeuProTTPalaThrArgSerHisLeuGlyArgArgLysCysSer 97  
DB 307 GGCTGCCCTTCCCTGGGCAACAAGAGCCACCTGGGCGAGAGAAAGTGACAGC 257

## RESULT 6

AAZ16609  
ID AAZ16609 standard; cDNA; 800 BP.  
XX  
AC AAZ16609;  
XX  
DT 12-OCT-1999 (first entry)  
XX  
DE Human gene expression product cDNA sequence SEQ ID NO:4079.  
XX  
KW Human; gene; gene expression product; diagnosis; therapy; probe; detection; mapping; tissue typing; profiling; forensic; cancer; genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9938972-A2.  
XX  
PD 05-AUG-1999.  
XX  
PF 28-JAN-1999; 99WO-US01619.  
XX  
PR 03-APR-1998; 98US-0080666.  
PR 28-JAN-1998; 98US-0072910.  
PR 24-FEB-1998; 98US-0075954.  
PR 31-MAR-1998; 98US-0080114.  
PR 03-APR-1998; 98US-0080515.  
XX  
PA (CHIR) CHIRON CORP.  
PA (HYSE-) HYSEQ INC.  
XX  
PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;  
PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;  
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;  
PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;  
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;  
XX  
DR WPI; 1999-494092/41.  
XX

Novel human genes and their expression products which are differentially expressed in different cell types

Claim 1; Page 1934; 2479pp; English.

The present invention describes a library of human polynucleotides comprising the sequences given in AAZ12532 to AAZ17779. Also described is a method of detecting differentially expressed genes correlated with the cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample from a cell



CC suspected of being cancerous, where the gene product is encoded by one  
CC of the 5248 polynucleotide sequences given in AA212532 to AA217779. The  
CC polynucleotides can be used as a source of primers and probes, which can  
CC be used for a variety of purpose, e.g. detection of expression levels,  
CC mapping, tissue typing or profiling, forensics, genetic analysis and  
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides  
CC can be used for raising antibodies for experimental, diagnostic and  
CC therapeutic purposes. The polynucleotides may also be used to construct  
CC arrays for diagnostics (which may be used to determine function of an  
CC encoded protein); and to detect differences in expression levels between  
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to  
CC identify a genetic predisposition or susceptibility to a disease such as  
CC cancer). The polynucleotides of the invention are especially used in the  
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,  
CC and lung cancer. The polynucleotides can also be used to screen for  
CC peptide analogues and antagonists.

XX  
SQ Sequence 800 BP; 199 A; 166 C; 202 G; 211 T; 22 other;

Alignment Scores:

Pred. No.:	2.59e-28	Length:	800
Score:	37.00	Matches:	37
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	38.14%	Indels:	0
DB:	20	Gaps:	0

US-09-854-133-586 (1-97) x AA216609 (1-800)

QY 61 GluserleuLeuCysProProSerProlysgluValThrcysArggluMetleuThrgly 80  
|||||  
Db 344 GAAAGCCTGTGTGTCCACCACTCCAAAGAGAGTTACCTGCAGGAAATGTTAACGGGA 403  
|||||

QY 81 GlyCysLeuProTPrAlaThrArgSerHisleuGlyArgArglyCysSer 97  
|||||  
Db 404 GGCTGCCTTCCCTGGGCAACAAGAGCCACCTGGGCGAGAGAAAGTGCAGC 454  
|||||

RESULT 7  
ABA09201  
ID ABA09201 standard; cDNA; 1958 BP.

XX  
AC ABA09201;

XX  
DT 11-JAN-2002 (first entry)

XX  
DE Human cystine/Glu transporter homologue cDNA, SEQ ID NO:977.

XX  
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;  
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;  
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
KW chronic inflammatory condition; proliferative retinopathy;  
KW atherosclerosis; coronary heart disease; arterial ischaemia;  
KW bone disorder; osteoporosis; vascular growth disorder;  
KW tissue regeneration; wound healing; infection; immune disorder;  
KW cell culture; drug screening; gene therapy; antiinflammatory;  
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;  
KW cytosstatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;  
KW antifungal; vulnerary; antilucer; ss.

XX  
OS Homo sapiens.

XX  
PN WO200157188-A2.

XX  
PD 09-AUG-2001.

XX  
PF 05-FEB-2001; 2001WO-US03800.

XX  
PR 03-FEB-2000; 2000US-0496914.

XX  
PR 27-APR-2000; 2000US-0560875.

XX  
PA (HYSE-) HYSEQ INC.

XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
XX WPI: 2001-457740/49.  
DR P-PSDB; ABB11957.

XX  
PT Human proteins and DNA encoding sequences useful for preventing,  
PT treating or ameliorating a medical condition in a mammalian subject  
PT e.g. arthritis and cancer

XX  
PS Claim 1; Page 835; 1963pp; English.

XX  
CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
CC invention also relates to vectors and recombinant host cells comprising a  
CC nucleotide of the invention, methods of producing the novel polypeptides,  
CC antibodies against the polypeptides, methods of detecting the nucleotides  
CC or polypeptides in a sample, and methods of identifying compounds which  
CC bind to polypeptides of the invention. Although novel, many of the  
CC polypeptides of the invention have homology to known proteins, thereby  
CC giving an insight into their probable biological activities, and hence  
CC potential therapeutic applications. The polypeptides of the invention may  
CC have various activities, including cytokine, cell proliferation or cell  
CC differentiation activities; stem cell growth factor activity;  
CC haematopoiesis regulatory activity; tissue growth factor activity;  
CC immunomodulatory activity; activin- or inhibin-related activities;  
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
CC thrombolytic activities; receptor or ligand activities; or may be  
CC involved in oncogenesis, cancer cell proliferation or metastasis.  
CC Depending on their biological activities, polypeptides and nucleotides of  
CC the invention are useful for preventing, treating or ameliorating medical  
CC conditions, e.g., by protein or gene therapy. Such conditions include  
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
CC vascular growth. Polypeptides involved with tissue regeneration and  
CC repair (or nucleic acids encoding them) may be used to promote wound  
CC healing (e.g., of burns, incisions and ulcers), while those with  
CC immunomodulatory activities may be used in the treatment of viral,  
CC bacterial and fungal infections in addition to immune disorders.  
CC Polypeptides with growth factor activity may be used in cell cultures to  
CC promote cell growth. For example, such polypeptides may be used to  
CC manipulate stem cells in culture to give rise to neuroepithelial cells  
CC that can be used to augment or replace cells damaged by illness,  
CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
CC may also be used in the diagnosis of the above conditions, and in drug  
CC screening techniques. The present sequence represents a cDNA encoding a  
CC novel human polypeptide of the invention.

XX  
SQ Sequence 1958 BP; 499 A; 418 C; 424 G; 617 T; 0 other;

Alignment Scores:

Pred. No.:	5.97e-28	Length:	1958
Score:	37.00	Matches:	37
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	38.14%	Indels:	0
DB:	22	Gaps:	0

US-09-854-133-586 (1-97) x ABA09201 (1-1958)

QY 61 GluserleuLeuCysProProSerProlysgluValThrcysArggluMetleuThrgly 80  
|||||  
Db 70 GAAAGCCTGTGTGTCCACCACTCCAAAGAGAGTTACCTGCAGGAAATGTTAACGGGA 129  
|||||

QY 81 GlyCysLeuProTPrAlaThrArgSerHisleuGlyArgArglyCysSer 97  
|||||  
Db 130 GGCTGCCTTCCCTGGGCAACAAGAGCCACCTGGGCGAGAGAAAGTGCAGC 180  
|||||

RESULT 8  
AAH98573  
ID AAH98573 standard; cDNA; 1958 BP.



PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-024617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.

PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
DR WPI; 2001-483426/52.  
XX  
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -  
XX  
PS Disclosure; SEQ ID NO 23451; 3071pp + Sequence Listing; English.  
XX  
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 123 BP; 32 A; 31 C; 37 G; 23 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 4.7e-28 Length: 123  
Score: 36.00 Matches: 36  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 37.11% Indels: 0  
DB: Gaps: 0  
  
US-09-854-133-586 (1-97) x AAK68639 (1-123)  
QY 61 GluSerLeuLeuCysProProSerProLySGluValThrCysArgGluMetLeuThrGly 80  
Db 14 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGGTTACCTGCAGGGAATGTTAACGGGA 73  
QY 81 GLYCysLeuProTPrAlaThrArgSerHisLeuGlyArgArgLysCys 96  
Db 74 GGCTGCCCTTCCCTGGGCAACAAGAGCCACCTGGGCAGAGAAAGTGC 121  
  
RESULT 10  
AAL37261  
ID AAL37261 standard; DNA; 29607 BP.  
XX  
AC AAL37261;  
XX  
DT 08-JAN-2002 (first entry)  
XX  
DE Human musculoskeletal system related polynucleotide SEQ ID NO 3626.  
XX  
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;  
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein;  
KW musculoskeletal system; ds.  
XX

OS Homo sapiens.  
XX  
PN WO200155367-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01338.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 18-AUG-2000; 2000US-0225759.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 23-AUG-2000; 2000US-0227182.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 06-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 12-SEP-2000; 2000US-0232081.  
PR 14-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.

PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0251990.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.



XX PI Rosen CA, Barash SC, Ruben SM;  
XX XX WPI; 2001-451937/48.  
XX XX  
PT Isolated polypeptide for treating, preventing and/ or prognosing  
PT disorders related to the musculoskeletal system including  
PT musculoskeletal cancers and also for testing and detection e.g.  
PT diagnosis -  
XX XX  
PS Example 2; SEQ ID NO 3626; 781pp + Sequence listing; English.  
XX XX  
CC The invention relates to novel genes (AAL34669-AAL37666) and proteins  
CC (ABB03087-ABB04109) associated with the musculoskeletal system useful  
CC for preventing, treating or ameliorating medical conditions e.g. by  
CC protein or gene therapy. The genes are isolated from a range of human  
CC tissues disclosed in the specification. The nucleic acids, proteins,  
CC antibodies and (ant)agonists are useful in the diagnosis, treatment  
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and  
CC other cancers of the adrenal gland, bone, bone marrow, breast,  
CC gastrointestinal tract, liver, lung, or urogenital; (b) immune  
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound  
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;  
CC and (f) infectious diseases such as viral, bacterial, fungal and  
CC parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX XX  
SQ Sequence 29607 BP; 8754 A; 4875 C; 5555 G; 10423 T; 0 other;  
Alignment Scores:  
Pred. No.: 0.191 Length: 29607  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
DB: 22 Gaps: 0  
US-09-854-133-586 (1-97) x AAL37261 (1-29607)  
QY 8 HisAlaserleuGlyAspserGluThrleuSerGln 19  
Db 24558 CACGCTAGCCTGGGTGACAGTGAACCTGTCTCAA 24593  
RESULT 11  
AAS36709  
ID AAS36709 standard; DNA; 32190 BP.  
XX XX  
AC AAS36709;  
XX XX  
DT 17-DEC-2001 (first entry)  
XX XX  
DE Human cardiovascular system antigen genomic DNA SEQ ID NO 2209.  
XX XX  
KW Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;  
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;  
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;  
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;  
KW ophthalmological; vulnerrary; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;  
KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; tissue regeneration;  
KW anti-infertility.  
XX XX  
OS Homo sapiens.  
XX XX  
PN WO200155321-A2.

XX PD 02-AUG-2001.  
XX XX  
PF 17-JAN-2001; 2001WO-US01340.  
XX XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX

DR WPI; 2001-451930/48.  
XX  
PT New cardiovascular system related polynucleotides and polypeptides,  
PT useful for diagnosing, treating and/or preventing disorders of the  
PT cardiovascular system -  
XX  
PS Claim 1; SEQ ID NO 2209; 674pp; English.  
XX  
CC Sequences AAS35741-AAS36942 represent genomic DNA molecules, which encode  
CC the cardiovascular system antigen polypeptides of the invention.  
CC Cardiovascular system antigens and their associated polynucleotides are  
CC useful in the diagnosis, treatment and prevention of various types of  
CC disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
CC chickens or sheep. A pathological condition can be determined by  
CC detecting the presence or absence of a mutation in a cardiovascular  
CC system antigen polynucleotide. The treatable disorders include autoimmune  
CC diseases such as rheumatoid arthritis, hyperproliferative disorders such  
CC as neoplasms of the breast or liver, cardiovascular disorders such as  
CC cardiac arrest, cerebrovascular disorders such as cerebral ischaemia,  
CC nervous system disorders such as Alzheimer's disease, infections caused  
CC by bacteria, viruses and fungi, ocular disorders such as corneal  
CC infection, endocrine disorders such as premature labour and infertility,  
CC gastrointestinal disorders such as Crohn's disease, renal disorders such  
CC as glomerulonephritis and respiratory disorders such as asthma and  
CC pleurisy. The polypeptides can also be used to aid wound healing, to  
CC prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, to regenerate tissues and in chemotaxis.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX

Alignment Scores:  
Pred. No.: 2.15 Length: 32190  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 22 Gaps: 0

US-09-854-133-586 (1-97) x AAS36709 (1-32190)

Oy 9 AlaserleuglyaspsergluThrleuserglu 19  
Db 30055 GCCAGCCTGGCGACGACGACTCTGTCTCAA 30087

RESULT 12

AAK82863  
ID AAK82863 standard; DNA; 136 BP.

XX AAK82863;

DT 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:37675.

DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX cytostatic; gene therapy; vaccine; metastasis; ds.

OS Homo sapiens.

XX WO200157182-A2.

PN 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US01354.

PF 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 11-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 14-JUL-2000; 2000US-0217496.  
PR 26-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 14-AUG-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 05-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 06-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.

PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249219.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 05-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 06-DEC-2000; 2000US-0256719.  
PR 08-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
  
(HUMA-) HUMAN GENOME SCI INC.  
Rosen CA, Barash SC, Ruben SM;  
WPI; 2001-483426/52.  
  
Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
useful for preventing, diagnosing and/or treating cancers and  
metastasis -  
Disclosure; SEQ ID NO 37675; 3071pp + Sequence listing; English.  
AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
amino acid sequences given in AAM82170 to AAM91921. (I) have cytosstatic  
activity, and can be used in gene therapy and vaccine production. (I)

CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64/03  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention.  
XX  
SQ Sequence 136 BP; 37 A; 34 C; 41 G; 24 T; 0 other;

Alignment Scores:

Pred. No.:	0.137	Length:	136
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	22	Gaps:	0

US-09-854-133-586 (1-97) x AAK82863 (1-136)

QY 10 SerLeuGLyAspSerGIuThrLeuSerGln 19  
Db 99 AGCCTGGGTGACAGTACAGTCTGTCTCAA 128  
|||||

RESULT 13

AAK82864  
ID AAK82864 standard; DNA; 136 BP.

XX AAK82864;

XX 07-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:37676.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.

PN WO200157182-A2.

PD 09-AUG-2001.

PF 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.



PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
DR WPI; 2001-483426/52.  
XX  
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating cancers and  
PT metastasis -  
XX  
PS Disclosure; SEQ ID NO 37676; 3071bp + Sequence Listing; English.  
XX  
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patients own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention.  
XX

SO Sequence 136 BP; 37 A; 34 C; 41 G; 24 T; 0 other;  
Alignment Scores:  
Pred. No.: 0.137 Length: 136  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 22 Gaps: 0  
US-09-854-133-586 (1-97) x AAK82864 (1-136)  
QY 10 SerLeuGIYaspSerGIuThrLeuSerGIu 19  
Db 99 AGCCTGGGTGACAGTGAGACTCTGTCTCAA 128  
RESULT 14  
AAC23449/C  
ID AAC23449 standard; cDNA: 207 BP.  
XX  
AC AAC23449;  
XX  
DT 06-OCT-2000 (first entry)  
XX  
DE Human secreted protein 5' EST, SEQ ID NO: 27524.  
XX  
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX  
OS Homo sapiens.  
XX  
PN EP1033401-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 21-FEB-2000; 2000EP-0200610.  
XX  
PR 26-FEB-1999; 99US-0122487.  
XX  
PA (GEST ) GENSET.  
XX  
PI Dumas Milne Edwards J, Duclert A, Giordano J;  
XX  
DR WPI; 2000-500381/45.  
XX  
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX  
PS Claim 1; SEQ ID 27524; 71pp + CD-ROM; English.  
XX  
CC The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
CC identified within the present sequence. The 5' ESTs were prepared from  
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
CC derived from the 5' ends of mRNAs and even in those cases where longer  
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.  
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC They are used to obtain upstream regulatory sequences and to design  
CC expression and secretion vectors.  
XX  
SQ Sequence 207 BP; 46 A; 44 C; 53 G; 60 T; 4 other;  
Alignment Scores:  
Pred. No.: 0.202 Length: 207  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 10.31% Indels: 0  
DB: 21 Gaps: 0

US-09-854-133-586 (1-97) x AAC23449 (1-207)

QY 10 SerLeuGlyAspSerGluThrLeuSerGln 19  
DB 102 AGCCTGGGTGACAGTGAACCCCTGTCTCAA 73

RESULT 15  
AAAA4648/C  
ID AAA44648 standard; cDNA; 291 BP.

AC AAA44648;  
XX  
DT 21-AUG-2000 (first entry)  
XX  
DE Human secreted expressed sequence tag SEQ ID NO:1223.  
XX  
KW Human; mouse; chicken; rat; secreted expressed sequence tag; SEST;  
KW expressed sequence tag; EST; probe; chemotactic; proliferative;  
KW immunomodulatory; haematopoietic; chemokinetic; analgesic; haemostatic;  
KW thrombolytic; antinflammatory; cytosolic; antibacterial; antifungal;  
KW antiviral; antidiabetic; antiasthmatic; vulnery; antiparkinsonian;  
KW antitumor; osteoprotective; neuroprotective; nootropic; antipsoriatic;  
KW cerebroprotective; anticonvulsant; antidepressant; gene therapy;  
KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;  
KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
KW central nervous system disorder; Alzheimer's disease; stroke;  
KW Parkinson's disease; Huntington's disease; coagulation disorder;  
KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
KW tumour; infection; depression; psoriasis; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200021991-A1.  
XX  
PD 20-APR-2000.  
XX  
PF 15-OCT-1999; 99WO-US24206.  
XX  
PR 15-OCT-1998; 98US-0104436.  
XX  
PA (GEMY ) GENETICS INST INC.  
XX  
PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
PI Merberg D, Treacy M, Bowman MR;  
XX  
DR WPI; 2000-317938/27.  
XX  
PT Isolated polynucleotides, and encoded proteins, comprising secreted  
PT expressed sequence tags (SESTs), useful for treating various disorders  
PT such as autoimmune, infectious, and central nervous system disorders -  
XX  
PS Claim 1; Page 516; 803pp; English.

AAAA3426 to AAA45925 represent specifically claimed secreted expressed  
CC sequence tags (SESTs), isolated from human, mouse, chicken and rat  
CC tissue sources. The SESTs can have a range of activities depending on  
CC the tissues they were isolated from. The activities include:  
CC chemotactic; proliferative; immunomodulatory; haematopoietic;  
CC chemokinetic; analgesic; haemostatic; thrombolytic; antinflammatory;  
CC cytosolic; antibacterial; antifungal; antiviral; antidiabetic;  
CC antiasthmatic; vulnery; antitumor; osteoprotective; neuroprotective;  
CC nootropic; antiparkinsonian; antipsoriatic; cerebroprotective;  
CC anticonvulsant; and antidepressant. The SESTs can be used for gene  
CC therapy and in vaccines. The SESTs are useful as probes for the  
CC identification and isolation of full-length cDNAs and genomic DNA  
CC molecules which correspond to the SESTs. Proteins encoded by the SESTs  
CC are useful in assays for determining biological activity and raising  
CC antibodies. They may be useful for treatment of autoimmune disorders  
CC (multiple sclerosis, insulin dependent diabetes), allergic conditions

CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
CC osteoporosis, osteoarthritis, central nervous system disorders  
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
CC disease), tumours, bacterial, fungal or viral infections, depression and  
CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given  
CC in the exemplification of the present invention.  
XX  
SQ Sequence 291 BP; 59 A; 83 C; 57 G; 92 T; 0 other;

Alignment Scores:  
Pred. No.: 0.277 Length: 291  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 21 Gaps: 0

US-09-854-133-586 (1-97) x AAA44648 (1-291)

QY 10 SerLeuGlyAspSerGluThrLeuSerGln 19  
DB 171 AGCCTGGGTGACAGTGAACCCCTGTCTCAA 142

RESULT 16  
AAS37013  
ID AAS37013 standard; cDNA; 291 BP.  
XX  
AC AAS37013;  
XX  
DT 17-DEC-2001 (first entry)  
XX  
DE Novel human diagnostic and therapeutic gene #71.  
XX  
KW Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200166753-A2.  
XX  
PD 13-SEP-2001.  
XX  
PF 09-MAR-2001; 2001WO-US07787.  
XX  
PR 09-MAR-2000; 2000US-0188609.  
XX  
PA (CHTR ) CHIRON CORP.  
PA (HYSE-) HYSEQ INC.  
XX  
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;  
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Kassam A, Lamson G;  
PI Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;  
PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;  
XX  
DR WPI; 2001-530177/58.  
XX  
PT New polynucleotides and polypeptides, useful for diagnosis and  
PT treatment of breast, lung and colon cancer -  
XX  
PS Claim 1; Page 613; 1193pp; English.

The invention relates to new polynucleotides and polypeptides, useful for  
CC diagnosis and treatment of breast, lung and colon cancer. The sequences  
CC can be used in detecting differentially expressed genes correlated with a  
CC cancerous state of a mammalian cell, comprising detecting at least one  
CC differentially expressed gene product in a test sample derived from a  
CC cell suspected of being cancerous. They can also be used to inhibit  
CC tumour growth by modulating expression of a gene product. AAS36943-  
CC AAS39338 represent novel human diagnostic and therapeutic coding  
CC sequences of the invention.  
XX  
SQ Sequence 291 BP; 76 A; 61 C; 80 G; 74 T; 0 other;

Alignment Scores:

Pred. No.:	0.277	Length:	291
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	22	Gaps:	0

US-09-854-133-586 (1-97) x AAS37013 (1-291)

QY 10 SerLeuGIYAspSerGIuThrLeuSerGIu 19

Db 97 AGCCTGGGTGACAGAGACTCTATCTCAA 126

RESULT 17

AAS37176

ID AAS37176 standard; cDNA; 291 BP.

XX AAS37176;

DT 17-DEC-2001 (first entry)

DE Novel human diagnostic and therapeutic gene #234.

KW Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.

OS Homo sapiens.

PN WO200166753-A2.

PD 13-SEP-2001.

PF 09-MAR-2001; 2001WO-US07787.

PR 09-MAR-2000; 2000US-0188609.

PA (CHIR ) CHIRON CORP.

PA (HYSE-) HYSEQ INC.

PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;

PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Kassam A, Lamson G;

PI Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;

PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;

DR WPI; 2001-530177/58.

PT New polynucleotides and polypeptides, useful for diagnosis and

PT treatment of breast, lung and colon cancer -

PS Claim 1; page 655; 1193pp; English.

CC The invention relates to new polynucleotides and polypeptides, useful for  
CC diagnosis and treatment of breast, lung and colon cancer. The sequences  
CC can be used in detecting differentially expressed genes correlated with a  
CC cancerous state of a mammalian cell, comprising detecting at least one  
CC differentially expressed gene product in a test sample derived from a  
CC cell suspected of being cancerous. They can also be used to inhibit  
CC tumour growth by modulating expression of a gene product. AAS36943-  
CC AAS39338 represent novel human diagnostic and therapeutic coding  
CC sequences of the invention.

SQ Sequence 291 BP; 74 A; 69 C; 77 G; 71 T; 0 other;

Alignment Scores:

Pred. No.:	0.277	Length:	291
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	22	Gaps:	0

US-09-854-133-586 (1-97) x AAS37176 (1-291)

QY 10 SerLeuGIYAspSerGIuThrLeuSerGIu 19

Db 28 AGCCTGGGTGACAGAGACTCTATCTCAA 57

RESULT 18

AAK75253

ID AAK75253 standard; DNA; 299 BP.

XX AAK75253;

DT 07-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:30065.

KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

KW cytostatic; gene therapy; vaccine; metastasis; ds.

OS Homo sapiens.

PN WO200157182-A2.

PD 09-AUG-2001.

PF 17-JAN-2001; 2001WO-US01354.

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225265.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226868.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227182.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.





KW detection; mapping; tissue typing; profiling; forensic; cancer;  
KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9938972-A2.  
XX  
PD 05-AUG-1999.  
XX  
PF 28-JAN-1999; 99WO-US01619.  
XX  
PR 03-APR-1998; 98US-0080666.  
PR 28-JAN-1998; 98US-0072910.  
PR 24-FEB-1998; 98US-0075954.  
PR 31-MAR-1998; 98US-0080114.  
PR 03-APR-1998; 98US-0080515.  
XX  
PA (CHIR ) CHIRON CORP.  
PA (HYSE-) HYSEQ INC.  
XX  
PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;  
PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;  
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;  
PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;  
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;  
XX  
DR WPI; 1999-494092/41.  
XX  
PT Novel human genes and their expression products which are  
PT differentially expressed in different cell types  
XX  
PS Claim 1; Page 1031; 2479pp; English.  
XX  
CC The present invention describes a library of human polynucleotides  
CC comprising the sequences given in AA212532 to AA217779. Also described is  
CC a method of detecting differentially expressed genes correlated with the  
CC cancerous state of a mammalian cell, comprising detecting at least one  
CC differentially expressed gene product in a test sample from a cell  
CC suspected of being cancerous, where the gene product is encoded by one  
CC of the 5248 polynucleotide sequences given in AA212532 to AA217779. The  
CC polynucleotides can be used as a source of primers and probes, which can  
CC be used for a variety of purpose, e.g. detection of expression levels,  
CC mapping, tissue typing or profiling, forensics, genetic analysis and  
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides  
CC can be used for raising antibodies for experimental, diagnostic and  
CC therapeutic purposes. The polynucleotides may also be used to construct  
CC arrays for diagnostics (which may be used to determine function of an  
CC encoded protein); and to detect differences in expression levels between  
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to  
CC identify a genetic predisposition or susceptibility to a disease such as  
CC cancer). The polynucleotides of the invention are especially used in the  
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,  
CC and lung cancer. The polynucleotides can also be used to screen for  
CC peptide analogues and antagonists.  
XX  
SQ Sequence 300 BP; 86 A; 63 C; 77 G; 74 T; 0 other;  
XX  
Alignment Scores:  
Pred. No.: 0.285 Length: 300  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 20 Gaps: 0  
US-09-854-133-586 (1-97) x AA214169 (1-300)  
OY 61 GluSerLeuIleuGysProProSerProLys 70  
Db 270 GAAAGCCTGTGTGTCACCATCTCCAAAG 299  
RESULT 20  
AA213802/c

ID AA213802 standard; cDNA; 300 BP.  
XX  
AC AA213802;  
XX  
DT 12-OCT-1999 (first entry)  
XX  
DE Human gene expression product cDNA sequence SEQ ID NO:1271.  
XX  
KW Human; gene; gene expression product; diagnosis; therapy; probe;  
KW detection; mapping; tissue typing; profiling; forensic; cancer;  
KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9938972-A2.  
XX  
PD 05-AUG-1999.  
XX  
PF 28-JAN-1999; 99WO-US01619.  
XX  
PR 03-APR-1998; 98US-0080666.  
PR 28-JAN-1998; 98US-0072910.  
PR 24-FEB-1998; 98US-0075954.  
PR 31-MAR-1998; 98US-0080114.  
PR 03-APR-1998; 98US-0080515.  
XX  
PA (CHIR ) CHIRON CORP.  
PA (HYSE-) HYSEQ INC.  
XX  
PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;  
PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;  
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;  
PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;  
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;  
XX  
DR WPI; 1999-494092/41.  
XX  
PT Novel human genes and their expression products which are  
PT differentially expressed in different cell types  
XX  
PS Claim 1; Page 950; 2479pp; English.  
XX  
CC The present invention describes a library of human polynucleotides  
CC comprising the sequences given in AA212532 to AA217779. Also described is  
CC a method of detecting differentially expressed genes correlated with the  
CC cancerous state of a mammalian cell, comprising detecting at least one  
CC differentially expressed gene product in a test sample from a cell  
CC suspected of being cancerous, where the gene product is encoded by one  
CC of the 5248 polynucleotide sequences given in AA212532 to AA217779. The  
CC polynucleotides can be used as a source of primers and probes, which can  
CC be used for a variety of purpose, e.g. detection of expression levels,  
CC mapping, tissue typing or profiling, forensics, genetic analysis and  
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides  
CC can be used for raising antibodies for experimental, diagnostic and  
CC therapeutic purposes. The polynucleotides may also be used to construct  
CC arrays for diagnostics (which may be used to determine function of an  
CC encoded protein); and to detect differences in expression levels between  
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to  
CC identify a genetic predisposition or susceptibility to a disease such as  
CC cancer). The polynucleotides of the invention are especially used in the  
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,  
CC and lung cancer. The polynucleotides can also be used to screen for  
CC peptide analogues and antagonists.  
XX  
SQ Sequence 300 BP; 62 A; 83 C; 62 G; 93 T; 0 other;  
XX  
Alignment Scores:  
Pred. No.: 0.285 Length: 300  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 20 Gaps: 0

US-09-854-133-586 (1-97) x AA213802 (1-300)

QY 10 SerleuGlyAspSerGluThrLeuSerGln 19  
Pred. No.: |||||  
Score: 10.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 10.31%  
DB: 20

AA213190 standard; cDNA; 300 BP.

AA213190;

12-OCT-1999 (first entry)

Human gene expression product cDNA sequence SEQ ID NO:659.

Human; gene; gene expression product; diagnosis; therapy; probe; detection; mapping; tissue typing; profiling; forensic; cancer; genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.

Homo sapiens.

WO9938972-A2.

05-AUG-1999.

28-JAN-1999; 99WO-US01619.

03-APR-1998; 98US-0080666.  
28-JAN-1998; 98US-0072910.  
24-FEB-1998; 98US-0075954.  
31-MAR-1998; 98US-0080114.  
03-APR-1998; 98US-0080515.

(CHIR ) CHIRON CORP.  
(HYSE-) HYSEQ INC.

Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;  
Escobedo J, Garcia PD, Garcia V, Glese K, Innis MA;  
Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;  
Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;  
Stache-Crain B, Sudduth-Klinger J, Williams LT;  
WPI; 1999-494092/41.

Novel human genes and their expression products which are differentially expressed in different cell types

Claim 1; Page 813; 2479pp; English.

The present invention describes a library of human polynucleotides comprising the sequences given in AA212532 to AA217779. Also described is a method of detecting differentially expressed genes correlated with the cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample from a cell suspected of being cancerous, where the gene product is encoded by one of the 5248 polynucleotide sequences given in AA212532 to AA217779. The polynucleotides can be used as a source of primers and probes, which can be used for a variety of purpose, e.g. detection of expression levels, mapping, tissue typing or profiling, forensics, genetic analysis and detection of polymorphisms. Polypeptides encoded by the polynucleotides can be used for raising antibodies for experimental, diagnostic and therapeutic purposes. The polynucleotides may also be used to construct arrays for diagnostics (which may be used to determine function of an encoded protein); and to detect differences in expression levels between two cells (e.g. to identify abnormal or diseased tissue in a human, to identify a genetic predisposition or susceptibility to a disease such as cancer). The polynucleotides of the invention are especially used in the diagnosis, prognosis and management of colorectal cancer, breast cancer, and lung cancer. The polynucleotides can also be used to screen for peptide analogues and antagonists.

Sequence 300 BP; 74 A; 72 C; 86 G; 68 T; 0 other;

Alignment Scores:

Pred. No.: 0.285  
Score: 10.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 10.31%  
DB: 20

US-09-854-133-586 (1-97) x AA213190 (1-300)

QY 10 SerleuGlyAspSerGluThrLeuSerGln 19  
Pred. No.: |||||  
Score: 10.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 10.31%  
DB: 20

AA337025 standard; cDNA; 303 BP.

17-DEC-2001 (first entry)

Novel human diagnostic and therapeutic gene #83.

Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.

Homo sapiens.

WO200166753-A2.

13-SEP-2001.

09-MAR-2001; 2001WO-US07787.

09-MAR-2000; 2000US-0188609.

(CHIR ) CHIRON CORP.  
(HYSE-) HYSEQ INC.

Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;  
Reinhard C, Randazzo F, Kennedy GC, Pot D, Kassam A, Lamson G;  
Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;  
Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;  
WPI; 2001-530177/58.

New polynucleotides and polypeptides, useful for diagnosis and treatment of breast, lung and colon cancer .

Claim 1; Page 616; 1193pp; English.

The invention relates to new polynucleotides and polypeptides, useful for diagnosis and treatment of breast, lung and colon cancer. The sequences can be used in detecting differentially expressed genes correlated with a cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample derived from a cell suspected of being cancerous. They can also be used to inhibit tumour growth by modulating expression of a gene product. AA336943-AA339338 represent novel human diagnostic and therapeutic coding sequences of the invention.

Sequence 303 BP; 79 A; 65 C; 84 G; 75 T; 0 other;

Alignment Scores:

Pred. No.: 0.288  
Score: 10.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 10.31%  
DB: 22

Length: 303  
Matches: 10  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x AAS37025 (1-303)  
OY 10 SerLeuGlyAspSergIuThrLeuSergIu 19  
Db 97 AGCCTGGGTGACGACGAGACTCTATCTCAA 126  
RESULT 23  
AAK75254  
ID AAK75254 standard; DNA; 310 BP.  
XX AC AAK75254;  
XX DT 07-NOV-2001 (first entry)  
XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:30066.  
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX KW cytostatic; gene therapy; vaccine; metastasis; ds.  
OS Homo sapiens.  
XX PN WO200157182-A2.  
XX PD 09-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US01354.  
XX PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.

```
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/haematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
PS Disclosure; SEQ ID NO 30066; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 310 BP; 97 A; 75 C; 90 G; 48 T; 0 other;
XX
Alignment Scores:
Pred. No.: 0.294 Length: 310
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.31% Indels: 0
DB: 22 Gaps: 0
XX
US-09-854-133-586 (1-97) x AAK75254 (1-310)
OY 10 SerleuglyaspserglutThrleusergin 19
DB 252 AGCCTGGCGACAGTGTCTCTCA 281
XX
RESULT 24
AAI92003
ID AAI92003 standard; cDNA; 343 BP.
XX
AC AAI92003;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 12063.
```

```
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
PR 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-514838/56.
XX P-PSDB; AAO12072.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
XX diagnosing and treating e.g. leukaemia, inflammation and immune
XX disorders -
XX
PS Claim 1; SEQ ID NO 12063; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 343 BP; 85 A; 87 C; 90 G; 81 T; 0 other;
XX
Alignment Scores:
Pred. No.: 0.323 Length: 343
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.31% Indels: 0
DB: 22 Gaps: 0
XX
US-09-854-133-586 (1-97) x AAI92003 (1-343)
OY 10 SerleuglyaspserglutThrleusergin 19
DB 94 AGCCTGGCGACAGCGAGACTCTGTCCAA 123
XX
RESULT 25
ABL63133/C
ID ABL63133 standard; DNA; 351 BP.
XX
XX ABL63133;
XX
DT 15-MAY-2002 (first entry)
XX
DE Breast cancer related gene sequence SEQ ID NO:1470.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
```



KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX  
OS Homo sapiens.  
XX WO200194629-A2.  
XX  
PD 13-DEC-2001.  
XX  
PF 30-MAY-2001; 2001WO-US10838.  
XX  
PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.  
PR 25-SEP-2000; 2000US-234923P.  
PR 25-SEP-2000; 2000US-234924P.  
PR 25-SEP-2000; 2000US-235077P.  
PR 25-SEP-2000; 2000US-235082P.  
PR 25-SEP-2000; 2000US-235134P.  
PR 25-SEP-2000; 2000US-235280P.  
PR 26-SEP-2000; 2000US-235637P.  
PR 26-SEP-2000; 2000US-235638P.  
PR 27-SEP-2000; 2000US-235711P.  
PR 27-SEP-2000; 2000US-235720P.  
PR 27-SEP-2000; 2000US-235840P.  
PR 27-SEP-2000; 2000US-235863P.  
PR 28-SEP-2000; 2000US-236028P.  
PR 28-SEP-2000; 2000US-236032P.  
PR 28-SEP-2000; 2000US-236033P.  
PR 28-SEP-2000; 2000US-236034P.  
PR 28-SEP-2000; 2000US-236109P.  
PR 28-SEP-2000; 2000US-236111P.  
PR 29-SEP-2000; 2000US-236842P.  
PR 29-SEP-2000; 2000US-236891P.  
PR 02-OCT-2000; 2000US-237172P.  
PR 02-OCT-2000; 2000US-237173P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237294P.  
PR 02-OCT-2000; 2000US-237295P.  
PR 02-OCT-2000; 2000US-237316P.  
PR 03-OCT-2000; 2000US-237425P.  
PR 03-OCT-2000; 2000US-237598P.  
PR 03-OCT-2000; 2000US-237604P.  
PR 03-OCT-2000; 2000US-237606P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 01-NOV-2000; 2000US-244867P.  
PR 01-NOV-2000; 2000US-245084P.  
XX  
PA (AVAL-) AVALON PHARM.  
XX  
PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Scoppet DR, Weaver Z;  
XX  
DR WPI; 2002-188264/24.  
XX  
PT Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
PT determining a change in expression of a gene of a signature gene set -  
XX  
PS Claim 1; SEQ ID 1470; 44pp; English.  
XX  
CC The present invention describes a method (M1) for screening for an  
CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in

CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilm's tumour.  
XX  
SQ Sequence 351 BP; 88 A; 92 C; 67 G; 101 T; 3 other;  
  
Alignment Scores:  
Pred. No.: 0.33 Length: 351  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 24 Gaps: 0  
  
US-09-854-133-586 (1-97) x ABL63133 (1-351)  
QY 10 SerLeuGlyAspSerGluThrLeuSergln 19  
Db 153 AGCCTGGGTGACAGCAGACCCCTGTCGCA 124  
  
RESULT 26  
ABV19284  
ID ABV19284 standard; cDNA; 371 BP.  
XX  
AC ABV19284;  
XX  
DT 13-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 19275.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PE 20-FEB-2001; 2001WO-US05171.  
XX  
PR 17-FEB-2000; 2000US-183319P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
DR WPI; 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer -  
XX  
PS Claim 1; Page 3158; 11750pp; English.  
XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 371 BP; 112 A; 87 C; 87 G; 85 T; 0 other;

Alignment Scores:  
Pred. No.: 0.348 Length: 371  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 23 Gaps: 0

US-09-854-133-586 (1-97) x ABV19284 (1-371)

OY 10 SerleuGlyAspSerGluThrLeuSergln 19  
|||||  
Db 223 AGCCTGGCGACAGTGAGACCTGTCTCAA 252

RESULT 27  
ABV13690  
ID ABV13690 standard; cDNA; 376 BP.  
XX  
AC ABV13690;  
XX  
DT 13-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 13681.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US05171.  
XX  
PR 17-FEB-2000; 2000US-183319P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
DR WPI; 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer -  
XX  
PS Claim 1; Page 2271-2272; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate

CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 376 BP; 102 A; 68 C; 103 G; 103 T; 0 other;

Alignment Scores:  
Pred. No.: 0.352 Length: 376  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 23 Gaps: 0

US-09-854-133-586 (1-97) x ABV13690 (1-376)

OY 10 SerleuGlyAspSerGluThrLeuSergln 19  
|||||  
Db 57 AGCCTGGCGACAGTGAGACTTGTCTCAA 86

RESULT 28  
AAI89561  
ID AAI89561 standard; cDNA; 394 BP.  
XX  
AC AAI89561;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 9621.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200164835-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 26-FEB-2001; 2001WO-US04927.  
XX  
PR 28-FEB-2000; 2000US-0515126.  
PR 18-MAY-2000; 2000US-0577409.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-514838/56.  
DR P-PSDB; AAO09630.  
XX  
PT Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -  
XX  
PS Claim 1; SEQ ID NO 9621; 1399pp + Sequence Listing; English.

CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX  
SQ Sequence 394 BP; 128 A; 69 C; 115 G; 82 T; 0 other;

## Alignment Scores:

Pred. No.:	0.368	Length:	394
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	22	Gaps:	0

US-09-854-133-586 (1-97) x AA189561 (1-394)

OY 10 SerLeuGLyAspSerGLuThrLeuSergln 19  
|||||

Db 117 AGCCTGGCGACAGTGAGACCTGTCTCAA 146

## RESULT 29

ABV04521  
ID ABV04521 standard; cDNA; 401 BP.

XX  
AC ABV04521;

XX  
DT 13-SEP-2002 (first entry)

XX  
DE Human prostate expression marker cDNA 4512.

XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX  
KW pharmacogenomic marker; gene; ss.

XX  
OS Homo sapiens.

XX  
PN WO200160860-A2.

XX  
PD 23-AUG-2001.

XX  
PF 20-FEB-2001; 2001WO-US05171.

XX  
PR 17-FEB-2000; 2000US-183319P.

XX  
PR 16-MAR-2000; 2000US-189862P.

XX  
PR 25-MAY-2000; 2000US-207454P.

XX  
PR 09-JUN-2000; 2000US-211314P.

XX  
PR 18-JUL-2000; 2000US-219007P.

XX  
PR 13-DEC-2000; 2000US-255281P.

XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX  
PI Schlegel R, Endege WO, Monahan JE;

XX  
DR WPI; 2001-662795/76.

XX  
PS Claim 1; Page 783; 11750pp; English.

XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;

CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate

CC cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer

CC in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound;

CC (g) determining whether prostate cancer has metastasized in a patient;

CC (h) assessing the aggressiveness or indolence of prostate cancer in a

CC patient;

CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX  
SQ Sequence 401 BP; 102 A; 77 C; 115 G; 104 T; 3 other;

## Alignment Scores:

Pred. No.:	0.374	Length:	401
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	23	Gaps:	0

US-09-854-133-586 (1-97) x ABV04521 (1-401)

OY 10 SerLeuGLyAspSerGLuThrLeuSergln 19  
|||||

Db 81 AGCCTGGCGACAGTGAGACCTGTCTCAA 110

## RESULT 30

ABV34803  
ID ABV34803 standard; cDNA; 401 BP.

XX  
AC ABV34803;

XX  
DT 16-SEP-2002 (first entry)

XX  
DE Human prostate expression marker cDNA 34794.

XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX  
KW pharmacogenomic marker; gene; ss.

XX  
OS Homo sapiens.

XX  
PN WO200160860-A2.

XX  
PD 23-AUG-2001.

XX  
PF 20-FEB-2001; 2001WO-US05171.

XX  
PR 17-FEB-2000; 2000US-183319P.

XX  
PR 16-MAR-2000; 2000US-189862P.

XX  
PR 25-MAY-2000; 2000US-207454P.

XX  
PR 09-JUN-2000; 2000US-211314P.

XX  
PR 18-JUL-2000; 2000US-219007P.

XX  
PR 13-DEC-2000; 2000US-255281P.

XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX  
PI Schlegel R, Endege WO, Monahan JE;

XX  
DR WPI; 2001-662795/76.

XX  
PS Claim 1; Page 7283; 11750pp; English.

XX  
CC The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;

CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate

CC cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer

CC in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound;

CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 401 BP; 104 A; 78 C; 114 G; 105 T; 0 other;

## Alignment Scores:

Pred. No.:	0.374	Length:	401
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	23	Gaps:	0

US-09-854-133-586 (1-97) x ABV34803 (1-401)

QY 10 SerLeuGlyAspSerGluThrLeuSergln 19  
|||||  
Db 82 AGCCTGGGCGACAGTGTCTCTCAA 111

## RESULT 31

ABV43654  
ID ABV43654 standard; cDNA; 401 BP.

AC ABV43654;

DT 16-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 43645.

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

OS Homo sapiens.

PN WO200160860-A2.

PD 23-AUG-2001.

PF 20-FEB-2001; 2001WO-US05171.

PR 17-FEB-2000; 2000US-183319P.

PR 16-MAR-2000; 2000US-189862P.

PR 25-MAY-2000; 2000US-207454P.

PR 09-JUN-2000; 2000US-211314P.

PR 18-JUL-2000; 2000US-219007P.

PR 13-DEC-2000; 2000US-255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

DR WPI; 2001-662795/76.

PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer  
XX  
PS Claim 1; Page 8690; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the  
CC specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;

CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 401 BP; 104 A; 78 C; 114 G; 105 T; 0 other;

## Alignment Scores:

Pred. No.:	0.374	Length:	401
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	23	Gaps:	0

US-09-854-133-586 (1-97) x ABV43654 (1-401)

QY 10 SerLeuGlyAspSerGluThrLeuSergln 19  
|||||  
Db 82 AGCCTGGGCGACAGTGTCTCTCAA 111

## RESULT 32

ABL84142/C  
ID ABL84142 standard; cDNA; 418 BP.

AC ABL84142;

DT 17-MAY-2002 (first entry)

DE Human ovarian cancer related cDNA clone SEQ ID NO:7120.

KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.

OS Homo sapiens.

PN WO200192581-A2.

PD 06-DEC-2001.

PF 29-MAY-2001; 2001WO-US17756.

PR 26-MAY-2000; 2000US-207484P.

PA (CORI-) CORIXA CORP.

PI Algate PA, Harlocker SL, Jones R;

DR WPI; 2002-122075/16.

PT Composition for therapy and diagnosis of ovarian cancer comprising  
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding  
PT polypeptide, antibody specific to polypeptide or T cell expressing  
PT polypeptide  
XX  
PS Claim 1; SEQ ID 7120; 489pp; English.

CC The present invention describes a composition (I) comprising: carriers  
CC and immunostimulants; and a polypeptide (II) of a ovarian tumour  
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence  
CC (S1) from the 10912 nucleotide sequences as given in ABL77023 to  
CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell  
CC population of (II), or antigen presenting cells that express (II).  
CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to  
CC (S1) can be used for detecting ovarian cancer in a patient's biological  
CC sample preferably serum or ovarian tissue. The method comprises  
CC contacting a biological sample from a patient with (IV), detecting the  
CC amount of polynucleotide hybridising to (IV) and comparing the amount to  
CC a predetermined cutoff value and thereby detecting ovarian cancer in the  
CC patient, where the amount of polynucleotide hybridising to (IV) is  
CC detected preferably by polymerase chain reaction (PCR). (I) comprising  
CC (III) and/or (II) is useful for stimulating and/or expanding T cells  
CC specific for an ovarian tumour protein comprising contacting T cells  
CC with (III) or (II). (III) is useful in design and preparation of  
CC ribozyme molecules for inhibiting expression of the tumour polypeptides



CC and proteins in tumour cells; and to isolate a full length gene from a  
CC suitable library e.g., a tumour cDNA library using well known  
CC techniques.  
XX  
SQ Sequence 418 BP; 93 A; 92 C; 77 G; 156 T; 0 other;  
Alignment Scores:  
Pred. No.: 0.389 Length: 418  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 24 Gaps: 0  
US-09-854-133-586 (1-97) x ABL84142 (1-418)  
QY 10 SerLeuGLYAspSerGIuThrLeuSerGIu 19  
Db 336 AGCCTGGAGACAGACGAGACTCTGTCTCAA 307  
RESULT 33  
AAI92526 standard; cDNA; 430 BP.  
XX  
AC AAI92526;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 12586.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorders; arthritis; inflammation; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200164835-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 26-FEB-2001; 2001WO-US04927.  
XX  
PR 28-FEB-2000; 2000US-0515126.  
PR 18-MAY-2000; 2000US-0577409.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-514838/56.  
DR P-PSDB; AAO12595.  
XX  
PT Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukaemia, inflammation and immune  
PT disorders -  
XX  
PS Claim 1; SEQ ID NO 12586; 1399pp + Sequence Listing; English.  
XX  
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX  
SQ Sequence 430 BP; 142 A; 78 C; 109 G; 101 T; 0 other;  
Alignment Scores:  
Pred. No.: 0.399 Length: 430  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 22 Gaps: 0  
US-09-854-133-586 (1-97) x AAI92526 (1-430)  
QY 10 SerLeuGLYAspSerGIuThrLeuSerGIu 19  
Db 104 AGCCTGGGTGACAGTGAGACCCGTGTCTCAA 133  
RESULT 34  
ABV49058 standard; cDNA; 440 BP.  
XX  
AC ABV49058;  
XX  
DT 17-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker CDNA 49049.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US05171.  
XX  
PR 17-FEB-2000; 2000US-183319P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
DR WPI; 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer -  
XX  
PS Claim 1; Page 9597; 11750pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX



PR	14-AUG-2000;	2000US-0225268.
PR	14-AUG-2000;	2000US-0225270.
PR	14-AUG-2000;	2000US-0225447.
PR	14-AUG-2000;	2000US-0225757.
PR	14-AUG-2000;	2000US-0225758.
PR	14-AUG-2000;	2000US-0225759.
PR	18-AUG-2000;	2000US-0226279.
PR	22-AUG-2000;	2000US-0226681.
PR	22-AUG-2000;	2000US-0226868.
PR	22-AUG-2000;	2000US-0227182.
PR	23-AUG-2000;	2000US-0227009.
PR	30-AUG-2000;	2000US-0228924.
PR	01-SEP-2000;	2000US-0229287.
PR	01-SEP-2000;	2000US-0229343.
PR	01-SEP-2000;	2000US-0229344.
PR	01-SEP-2000;	2000US-0229345.
PR	05-SEP-2000;	2000US-0229513.
PR	05-SEP-2000;	2000US-0230437.
PR	06-SEP-2000;	2000US-0230438.
PR	08-SEP-2000;	2000US-0231242.
PR	08-SEP-2000;	2000US-0231243.
PR	08-SEP-2000;	2000US-0231244.
PR	08-SEP-2000;	2000US-0231413.
PR	08-SEP-2000;	2000US-0231414.
PR	08-SEP-2000;	2000US-0232080.
PR	08-SEP-2000;	2000US-0232081.
PR	12-SEP-2000;	2000US-0231968.
PR	14-SEP-2000;	2000US-0232397.
PR	14-SEP-2000;	2000US-0232398.
PR	14-SEP-2000;	2000US-0232399.
PR	14-SEP-2000;	2000US-0232400.
PR	14-SEP-2000;	2000US-0233063.
PR	14-SEP-2000;	2000US-0233065.
PR	14-SEP-2000;	2000US-0233066.
PR	21-SEP-2000;	2000US-0234223.
PR	21-SEP-2000;	2000US-0234274.
PR	25-SEP-2000;	2000US-0234997.
PR	25-SEP-2000;	2000US-0234998.
PR	26-SEP-2000;	2000US-0235484.
PR	27-SEP-2000;	2000US-0235834.
PR	27-SEP-2000;	2000US-0235836.
PR	29-SEP-2000;	2000US-0236327.
PR	29-SEP-2000;	2000US-0236367.
PR	29-SEP-2000;	2000US-0236368.
PR	29-SEP-2000;	2000US-0236369.
PR	29-SEP-2000;	2000US-0236370.
PR	02-OCT-2000;	2000US-0236802.
PR	02-OCT-2000;	2000US-0237037.
PR	02-OCT-2000;	2000US-0237038.
PR	02-OCT-2000;	2000US-0237039.
PR	02-OCT-2000;	2000US-0237040.
PR	13-OCT-2000;	2000US-0239935.
PR	13-OCT-2000;	2000US-0239937.
PR	20-OCT-2000;	2000US-0240960.
PR	20-OCT-2000;	2000US-0241221.
PR	20-OCT-2000;	2000US-0241785.
PR	20-OCT-2000;	2000US-0241786.
PR	20-OCT-2000;	2000US-0241787.
PR	20-OCT-2000;	2000US-0241808.
PR	20-OCT-2000;	2000US-0241809.
PR	20-OCT-2000;	2000US-0244617.
PR	01-NOV-2000;	2000US-0246474.
PR	08-NOV-2000;	2000US-0246475.
PR	08-NOV-2000;	2000US-0246476.
PR	08-NOV-2000;	2000US-0246477.
PR	08-NOV-2000;	2000US-0246478.
PR	08-NOV-2000;	2000US-0246523.
PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.

```

PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

PA (HUMA-) HUMAN GENOME SCI INC.
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465569/50.
DR
XX
XX Isolated nucleic acid molecule encoding excretory system antigen is
PT used in preventing, treating or ameliorating a medical condition -
XX
XX Example 2; SEQ ID NO 715; 574pp + Sequence Listing; English.
PS
XX
XX The invention relates to novel excretory system related human
CC polynucleotides (AA198567-AA199503) and the encoded proteins
CC (AAM99594-AAM99913) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy, especially
CC disorders related to the excretory system. The genes are isolated
CC from a range of human tissues disclosed in the specification. The
CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 459 BP; 106 A; 108 C; 123 G; 122 T; 0 other;

```

Alignment Scores:

Pred. No.:	0.424	Length:	459
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	22	Gaps:	0

US-09-854-133-586 (1-97) x AAI98951 (1-459)

QY	10	SerLeuGlyAspSergluThrIeuSergln	19
DB	373	AGCCTGGGTGACAGCGAGACCTATCTCAA	344

RESULT 37  
AAI63301/C

ID AAI63301 standard; DNA; 459 BP.

XX AAI63301;

DT 22-OCT-2001 (first entry)

DE Human kidney related polynucleotide SEQ ID NO 616.

Human; kidney antigen; immunosuppressive; antiarthritic; antirheumatic;  
antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
neuroprotective; antibacterial; virucide; fungicide;  
ophthalmological; antiallergic; hepatotropic; antidiabetic;  
antiinflammatory; antitumor; anticonvulsant; antiparasitic;  
gene therapy; cancer; immune disorder; cardiovascular disorder;  
neurological disease; infection; ds.

OS Homo sapiens.

PN WO200155323-A2.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001WO-US01343.

XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235836.  
PR 27-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.





DB: 23 Gaps: 0

US-09-854-133-586 (1-97) x ABV49991 (1-486)

QY 10 SerLeuGIyAspSerGIuThrLeuSerGI n 19  
|||||  
Db 444 AGCCTGGGCGACAGCAGACTGTCTCA 473

RESULT 39

ABV46023  
ID ABV46023 standard; cDNA; 490 BP.  
XX  
AC ABV46023;  
XX  
DT 16-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 46014.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US05171.  
XX  
PR 17-FEB-2000; 2000US-183319P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
DR WPI; 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer  
XX  
PS Claim 1; Page 9092; 11750pp; English.  
XX

CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 490 BP; 155 A; 116 C; 107 G; 112 T; 0 other;

Alignment Scores:

Pred. No.:	0.451	Length:	490
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	23	Gaps:	0

US-09-854-133-586 (1-97) x ABV46023 (1-490)

QY 10 SerLeuGIyAspSerGIuThrLeuSerGI n 19  
|||||  
Db 283 AGCCTGGGTGACAGTGAACCTGTCTCA 312

RESULT 40

ABV49206  
ID ABV49206 standard; cDNA; 491 BP.  
XX  
AC ABV49206;  
XX  
DT 17-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 49197.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US05171.  
XX  
PR 17-FEB-2000; 2000US-183319P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
DR WPI; 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer  
XX  
PS Claim 1; Page 9620; 11750pp; English.  
XX

CC The invention relates to an isolated nucleic acid molecule (I) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (I) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 491 BP; 172 A; 82 C; 120 G; 117 T; 0 other;

Alignment Scores:

Pred. No.:	0.452	Length:	491
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	23	Gaps:	0

US-09-854-133-586 (1-97) x ABV49206 (1-491)

QY 10 SerLeuGlyAspSerGluThrLeuSerGln 19

Db 276 AGCCTGGGCGACGACGACTCTGTCTCAA 305

## RESULT 41

AAS30313/C

ID AAS30313 standard; cDNA; 497 BP.

XX AAS30313;

DT 21-NOV-2001 (first entry)

DE DNA encoding novel prostate gland antigen, Seq ID No 63.

XX Human; neurotropic; neuroprotective; cytosolic; antiparkinsonian;  
KW antianemic; dermatological; immunosuppressive; antiinflammatory;  
KW antiarthritic; antirheumatic; virucide; hepatotropic; nephrotropic;  
KW osteopathic; prostate gland; prostatitis; adenocarcinoma; hair loss;  
KW prostatic; malacoplakia; adenocarcinoma; benign prostatic hypertrophy;  
KW hyperplasia; carcinoma; prostate neoplastic disorder; skin aging;  
KW reproductive system disorder; autoimmune disorder; urinary system;  
KW systemic lupus erythematosus; rheumatoid arthritis; cardiovascular;  
KW blood-related disorder; hyperproliferative disorder; respiratory;  
KW neurological disorder; endocrine disorder; inflammatory disorder;  
KW liver disorder; wound healing; food preservative; ss.

XX Homo sapiens.

XX WO200155447-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01330.

XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.

PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX  
XX

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-476223/51.  
P-PSDB; AAU18872.

Novel isolated prostate gland related polypeptide useful for diagnosis and treatment of disorders of prostate such as prostatodystonia, prostatosis, prostatitis, benign prostatic hypertrophy and malacoplakia

Claim 1; SEQ ID NO 63; 512pp; English.

CC The invention relates to novel isolated prostate gland related nucleic acids (I) and polypeptides (II). (I) and (II) are useful for diagnosis, prognosis, prevention, and/or treatment of diseases and/or disorders, the prostate such as acute non-bacterial prostatitis, chronic non-bacterial prostatitis, acute bacterial prostatitis, prostatodystonia, prostatosis, granulomatous prostatitis, malacoplakia, benign prostatic hypertrophy or hyperplasia, and prostate neoplastic disorders, including adenocarcinomas, transitional cell carcinomas, ductal carcinomas, and squamous cell carcinomas. (I), (II) and antibody to (II) are useful for diagnosing and treating reproductive system disorders (Paget's disease), autoimmune disorders (systemic lupus erythematosus, rheumatoid arthritis), blood-related disorders (sickle cell anaemia), hyperproliferative disorders (urinary system disorders), (glomerulonephritis), cardiovascular disorders (arrhythmias), respiratory disorders, musculoskeletal system disorders, neural activity and neurological disorders (Alzheimer's disease and Parkinson's disease), endocrine disorders (Addison's disease), gastrointestinal disorders (inflammatory disorders), liver disorders (biliary liver cirrhosis), pancreatic and gall bladder disorders, disorders of the large intestine, developmental and inherited disorders, diseases at the cellular level, and wound healing and epithelial cell proliferation. (I) or (II) is

Alignment Scores:

Pred. No.: 0.457  
Score: 10.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 10.31%  
DB: 22

Length: 497  
Matches: 10  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x AAS30313 (1-497)

OY 10 SerLeuGlyAspSerCluThrLeuSerGln 19  
Db 184 AGCCTGGTGCAGTGCAGACCCCTGCTCAA 155  
RESULT 42  
AAL01994/c  
ID AAL01994 standard; cDNA; 497 BP.  
XX  
AC AAL01994;  
XX  
DT 21-NOV-2001 (first entry)  
XX  
DE Human reproductive system related antigen cDNA SEQ ID NO: 1995.  
XX  
KW Human; reproductive system related antigen; reproductive system disorder;  
XX cancer; gene therapy; ss.  
OS Homo sapiens.  
XX  
PN WO200155320-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01339.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225457.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.



PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 01-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.

PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465570/50.  
P-PSDB; AAM96024.

Isolated nucleic acid molecule encoding a reproductive system antigen  
is used in preventing, treating or ameliorating a medical condition -  
Claim 1; SEQ ID NO 1995; 1297pp + Sequence Listing; English.

The present invention provides the protein and coding sequences of a  
number of human reproductive system related antigens. These can be used  
in the prevention and treatment of reproductive system disorders,  
including cancer. The present sequence is a coding sequence of the  
invention.

Sequence 497 BP; 97 A; 136 C; 119 G; 140 T; 5 other;  
SQ

Alignment Scores: 0.457 Length: 497  
Pred. No.: 10.00 Matches: 10  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 10.31% Indels: 0  
Query Match: 22 Gaps: 0  
DB:

US-09-854-133-586 (1-97) x AAL01994 (1-497)

QY 10 SerleuglyaspSergluthrleusergln 19  
|||||

Db 184 AGCCTGGTGACAGTGAGACCCCTGTCTCAA 155

RESULT 43

AAH09380/c  
ID AAH09380 standard; cDNA; 509 BP.

XX AAH09380;

XX AC  
XX 26-JUN-2001 (first entry)

XX DE Human cDNA clone (3'-primer) SEQ ID NO:6215.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

XX OS Homo sapiens.

XX PN EP1074617-A2.

XX PD 07-FEB-2001.

XX PF 28-JUL-2000; 2000EP-0116126.

XX PR 29-JUL-1999; 99JP-0248036.

XX PR 27-AUG-1999; 99JP-0300253.

XX PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.  
XX 09-JUN-2000; 2000JP-0241899.  
PA (HELI-) HELIX RES INSE.  
XX  
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
DR WPI; 2001-318749/34.  
XX  
PT Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
XX full-length cDNAs.  
PS Claim 3; SEQ ID 6215; 2537pp + CD ROM; English.  
XX  
CC The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
XX of the present invention.  
SQ Sequence 509 BP; 108 A; 128 C; 131 G; 132 T; 10 other;  
Alignment Scores:  
Pred. No.:  
Score: 0.467  
Length: 509  
Percent Similarity: 10.00  
Matches: 10  
Best Local Similarity: 100.00%  
Conservative: 0  
Query Match: 100.00%  
Mismatch: 0  
DB: 10.31%  
Indels: 0  
Gaps: 0  
US-09-854-133-586 (1-97) x AAH09380 (1-509)  
QY 10 SerLeuGlyAspSerGluThrLeuSerGln 19  
DB 35 AGTCTGGGTGACAGCGAGACCCCTATCTCAA 6  
RESULT 44  
ABA59875/c  
ID ABA59875 standard; DNA; 520 BP.  
XX  
AC ABA59875;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Human foetal liver single exon nucleic acid probe #8180.  
XX  
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.  
OS Homo sapiens.  
XX  
PN WO200157277-A2.  
XX  
PD 09-AUG-2001.

XX  
PF 30-JAN-2001; 2001WO-US00669.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
PI WPI; 2001-483447/52.  
DR  
XX  
XX Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human fetal liver -  
PS Claim 1; SEQ ID NO 8180; 639pp + sequence listing; English.  
XX  
CC The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC fetal liver. The present sequence is a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 520 BP; 135 A; 123 C; 128 G; 134 T; 0 other;  
Alignment Scores:  
Pred. No.:  
Score: 0.477  
Length: 520  
Percent Similarity: 10.00  
Matches: 10  
Best Local Similarity: 100.00%  
Conservative: 0  
Query Match: 100.00%  
Mismatch: 0  
DB: 10.31%  
Indels: 0  
Gaps: 0  
US-09-854-133-586 (1-97) x ABA59875 (1-520)  
QY 26 GluArgLysLysLysArgGluArgLysPhe 35  
DB 262 GAAAGGAAAAAGAGAGAGAGAAATT 233  
RESULT 45  
AAK08145/c  
ID AAK08145 standard; DNA; 520 BP.  
XX  
AC AAK08145;  
XX  
DT 05-NOV-2001 (first entry)  
XX  
DE Human brain expressed single exon probe SEQ ID NO: 8136.  
XX  
KW Human; brain expressed exon; gene expression analysis; probe;  
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
KW epilepsy; cancer; ss.  
OS Homo sapiens.  
XX  
PN WO200157275-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00667.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.

XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX  
DR WPI; 2001-483446/52.

XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT brains -

XX  
PS Example 4; SEQ ID NO: 8136; 650pp + Sequence Listing; English.

XX  
CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is one of the probes of the  
CC invention.

XX  
SQ Sequence 520 BP; 135 A; 123 C; 128 G; 134 T; 0 other;

#### Alignment Scores:

Pred. No.:	0.477	Length:	520
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	22	Gaps:	0

US-09-854-133-586 (1-97) x AAK08145 (1-520)

QY 26 GluArgLysLysLysArgGluArgLysPhe 35  
|||  
DB 262 GAAAGGAAAAAAGAGAGAGAGAAAAATTT 233

Search completed: May 11, 2003, 17:16:00  
Job time : 219.735 secs





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 11, 2003, 16:16:15 ; Search time 54.0796 Seconds  
(without alignments)  
550.071 Million cell updates/sec

Title: US-09-854-133-586

Perfect score: 97

Sequence: 1 EVEVSRDHSALGDSFSLTSLQ.....LTGCLPWATFRSHLGRKCS 97

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338361 residues

Word size: 1

Total number of hits satisfying chosen parameters: 878589

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 60 summaries

Command line parameters:

-MODEL-frame+ p2n.model -DEV=xlh  
-Q-/cgn2\_1/USPTO\_spool/US09854133/runat\_05052003\_174133\_728/app\_query.fasta\_1.462  
-DB-Issued\_Patents\_NA -OFMT-fastap -SUFFIX=oligo.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX=oligo -TRANS-human40.cdi  
-LIST=60 -DOCALIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=45 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09854133 @CGN\_1\_1\_47 @runat\_05052003\_174133\_728 -NCPV=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	9	9.3	36	1	US-08-004-800-18
c 2	9	9.3	36	1	US-08-413-813-18
c 3	9	9.3	36	2	US-08-467-346-18
c 4	9	9.3	1753	4	US-08-205-697A-6
c 5	9	9.3	1753	4	US-08-702-525-6
c 6	9	9.3	1753	5	PCT-US95-02576-6
c 7	9	9.3	4092	4	US-09-306-595C-5
c 8	9	9.3	4136	4	US-09-103-875-2
c 9	9	9.3	12847	1	US-08-550-715-1
c 10	9	9.3	72928	3	US-09-009-913-1
c 11	9	9.3	152331	3	US-09-128-155-16
c 12	8	8.2	36	1	US-08-004-800-16

c 13	8	8.2	36	1	US-08-413-813-16	Sequence 16, Appl
c 14	8	8.2	36	2	US-08-467-346-16	Sequence 16, Appl
c 15	8	8.2	139	4	US-09-084-120-18	Sequence 18, Appl
c 16	8	8.2	291	1	US-07-922-723A-7	Sequence 7, Appl
c 17	8	8.2	291	1	US-07-799-828C-7	Sequence 7, Appl
c 18	8	8.2	291	1	US-08-074-275-7	Sequence 7, Appl
c 19	8	8.2	291	1	US-08-480-366-7	Sequence 7, Appl
c 20	8	8.2	291	1	US-07-952-277A-7	Sequence 7, Appl
c 21	8	8.2	299	2	US-08-623-906A-19	Sequence 19, Appl
c 22	8	8.2	350	2	US-08-623-906A-17	Sequence 17, Appl
c 23	8	8.2	350	3	US-09-157-177-110	Sequence 110, App
c 24	8	8.2	376	2	US-08-623-906A-18	Sequence 18, Appl
c 25	8	8.2	388	2	US-08-623-906A-13	Sequence 13, Appl
c 26	8	8.2	412	4	US-09-018-584A-22	Sequence 22, Appl
c 27	8	8.2	454	2	US-08-623-906A-6	Sequence 6, Appl
c 28	8	8.2	485	4	US-09-385-982-31	Sequence 31, Appl
c 29	8	8.2	494	3	US-08-332-766A-22	Sequence 22, Appl
c 30	8	8.2	1005	3	US-09-103-359-4	Sequence 4, Appl
c 31	8	8.2	1255	1	US-08-097-938-3	Sequence 3, Appl
c 32	8	8.2	1255	1	US-08-476-000-3	Sequence 3, Appl
c 33	8	8.2	1255	1	US-08-472-840-3	Sequence 3, Appl
c 34	8	8.2	1255	2	US-08-476-976-3	Sequence 3, Appl
c 35	8	8.2	1255	3	US-08-474-410-3	Sequence 3, Appl
c 36	8	8.2	1255	4	US-08-486-673B-3	Sequence 3, Appl
c 37	8	8.2	2186	4	US-09-184-001-1	Sequence 1, Appl
c 38	8	8.2	2310	3	US-08-458-434A-1	Sequence 1, Appl
c 39	8	8.2	2415	3	US-09-019-689-1	Sequence 1, Appl
c 40	8	8.2	2558	4	US-09-184-001-3	Sequence 3, Appl
c 41	8	8.2	3892	2	US-08-555-723B-3	Sequence 3, Appl
c 42	8	8.2	3892	3	US-09-123-465-3	Sequence 3, Appl
c 43	8	8.2	4192	4	US-09-122-126B-1	Sequence 1, Appl
c 44	8	8.2	9299	3	US-08-458-434A-7	Sequence 7, Appl
c 45	8	8.2	11461	3	US-08-669-161A-29	Sequence 29, Appl
c 46	8	8.2	11811	4	US-09-078-294-7	Sequence 7, Appl
c 47	8	8.2	28720	4	US-09-341-587-7	Sequence 7, Appl
c 48	8	8.2	45546	4	US-09-146-053-6	Sequence 6, Appl
c 49	8	8.2	45546	4	US-09-146-053-6	Sequence 6, Appl
c 50	8	8.2	45716	4	US-08-965-048-5	Sequence 5, Appl
c 51	8	8.2	45989	4	US-08-965-048-6	Sequence 6, Appl
c 52	8	8.2	59065	4	US-09-813-817-3	Sequence 3, Appl
c 53	8	8.2	59065	4	US-09-978-197-3	Sequence 3, Appl
c 54	8	8.2	62804	4	US-09-800-960-3	Sequence 3, Appl
c 55	8	8.2	176373	3	US-09-128-155-17	Sequence 17, Appl
c 56	7	7.2	26	2	US-08-291-011-8	Sequence 8, Appl
c 57	7	7.2	26	4	US-09-266-065-8	Sequence 8, Appl
c 58	7	7.2	36	2	US-08-291-011-4	Sequence 4, Appl
c 59	7	7.2	36	2	US-08-291-011-5	Sequence 5, Appl
c 60	7	7.2	36	4	US-09-266-065-4	Sequence 4, Appl

#### ALIGNMENTS

RESULT 1  
US-08-004-800-18/c  
Sequence 18, Application US/08004800  
Patent No. 5426180  
GENERAL INFORMATION:  
APPLICANT: KOOL, Eric T.  
TITLE OF INVENTION: SINGLE-STRANDED, CIRCULAR  
TITLE OF INVENTION: OLIGONUCLEOTIDES  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/004,800  
;; FILING DATE: 19930111  
;; CLASSIFICATION: 514  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: McNulty, William E.  
;; REGISTRATION NUMBER: 22,606  
;; REFERENCE/DOCKET NUMBER: 8085ZY  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (516) 742-4343  
;; TELEFAX: (516) 742-4366  
;; TELETYPE: 230 901 SANS UR  
;; INFORMATION FOR SEQ ID NO: 18:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 36 base pairs  
;; TYPE: NUCLEIC ACID  
;; STRANDEDNESS: single  
;; TOPOLOGY: circular  
;; MOLECULE TYPE: DNA  
US-08-004-800-18

Alignment Scores:  
Pred. No.: 0.0797 Length: 36  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.28% Indels: 0  
DB: 1 Gaps: 0

US-09-854-133-586 (1-97) x US-08-004-800-18 (1-36)

OY 26 GluArgLysLysLysArgGluArgLys 34  
Db 33 GAGAGAAAAAAGAGAGAGAGAAAA 7

RESULT 2  
US-08-413-813-18/c  
Sequence 18, Application US/08413813  
Patent No. 5683874  
GENERAL INFORMATION:  
APPLICANT: KOOL, Eric T.  
TITLE OF INVENTION: SINGLE-STRANDED, CIRCULAR OLIGONUCLEOTIDES  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/413,813  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Digilio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 8085ZYX  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELETYPE: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: circular

US-08-413-813-18

Alignment Scores:  
Pred. No.: 0.0797 Length: 36  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.28% Indels: 0  
DB: 1 Gaps: 0

US-09-854-133-586 (1-97) x US-08-413-813-18 (1-36)

OY 26 GluArgLysLysLysArgGluArgLys 34  
Db 33 GAGAGAAAAAAGAGAGAGAGAAAA 7

RESULT 3  
US-08-467-346-18/c  
Sequence 18, Application US/08467346  
Patent No. 5872105  
GENERAL INFORMATION:  
APPLICANT: KOOL, Eric T.  
TITLE OF INVENTION: SINGLE-STRANDED, CIRCULAR OLIGONUCLEOTIDES  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,346  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/413,813  
FILING DATE: 30-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Digilio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 8085ZYX  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELETYPE: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: circular  
US-08-467-346-18

Alignment Scores:  
Pred. No.: 0.0797 Length: 36  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.28% Indels: 0  
DB: 2 Gaps: 0

US-09-854-133-586 (1-97) x US-08-467-346-18 (1-36)

OY 26 GluArgLysLysLysArgGluArgLys 34  
Db 33 GAGAGAAAAAAGAGAGAGAGAAAA 7

RESULT 4  
US-08-205-697A-6/c  
; Sequence 6, Application US/08205697A  
; Patent No. 6218510  
; GENERAL INFORMATION:  
; APPLICANT: Sharpe, Arlene H.  
; APPLICANT: Borriello, Francescopaulo  
; APPLICANT: Freeman, Gordon J.  
; APPLICANT: Nadler, Lee M.  
; TITLE OF INVENTION: No. 6218510el Forms of T Cell Costimulatory Molecules  
; TITLE OF INVENTION: and Uses Therefor  
; NUMBER OF SEQUENCES: 61  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/205,697A  
; FILING DATE: 02-Mar-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragouras, Amy E.  
; REGISTRATION NUMBER: 36,207  
; REFERENCE/DOCKET NUMBER: BWI-120  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617)227-7400  
; TELEFAX: (617)227-5941  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1753 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; US-08-205-697A-6

Alignment Scores:  
Pred. No.: 2.96 Length: 1753  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.28% Indels: 0  
DB: 4 Gaps: 0

US-09-854-133-586 (1-97) x US-08-205-697A-6 (1-1753)

QY 23 ArgLysLysGluArgLysLysLysArg 31  
|||||  
Db 740 AGAAAAAGAGAGAGAAAAAGAA 714

RESULT 5  
US-08-702-525-6/c  
; Sequence 6, Application US/08702525  
; Patent No. 6294660  
; GENERAL INFORMATION:  
; APPLICANT: Sharpe, Sharpe  
; APPLICANT: Borriello, Francescopaulo  
; APPLICANT: Freeman, Gordon  
; APPLICANT: Nadler, Lee  
; TITLE OF INVENTION: No. 6294660el Forms of T Cell Costimulatory  
; TITLE OF INVENTION: Molecules and Uses Therefor  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 28 State Street  
; CITY: Boston

STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/702,525  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/205,697  
FILING DATE: 02-Mar-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: BWI-120CPUS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1753 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-702-525-6

Alignment Scores:  
Pred. No.: 2.96 Length: 1753  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.28% Indels: 0  
DB: 4 Gaps: 0

US-09-854-133-586 (1-97) x US-08-702-525-6 (1-1753)

QY 23 ArgLysLysGluArgLysLysLysArg 31  
|||||  
Db 740 AGAAAAAGAGAGAGAAAAAGAA 714

RESULT 6  
PCT-US95-02576-6/c  
; Sequence 6, Application PC/TUS9502576  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Novel Forms of T Cell Costimulatory Molecules  
; TITLE OF INVENTION: and Uses Therefor  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 State Street, suite 510  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/02576  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/205,697  
; FILING DATE: 02-Mar-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mandragouras, Amy E.  
; REGISTRATION NUMBER: 36,207





GENERAL INFORMATION:  
APPLICANT: Bowie, Lemuel J.  
TITLE OF INVENTION: Human  $\alpha$ -Thalassemia Mutations as a Predictor of  
TITLE OF INVENTION: Blood-Related Disorders  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/550,715  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Gass, David A.  
REGISTRATION NUMBER: 38,153  
REFERENCE/DOCKET NUMBER: 28493/32834  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12847 base pairs  
TYPE: nucleic acid.  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(6703..6797, 6915..7119, 7262..7387)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(10514..10608, 10726..10930, 11080..11205)  
US-08-550-715-1  
Alignment Scores:  
Pred. No.: 18.9 Length: 12847  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.28% Indels: 0  
Gaps: 0  
DB: 1  
US-09-854-133-586 (1-97) x US-08-550-715-1 (1-12847)  
QY 1 GluValGluValSerArgAspHisAla 9  
|||||  
Db 4510 GAGGTGAGGTGAGCGAGATCACGCC 4536  
RESULT 10  
US-09-009-913-1/c  
Sequence 1, Application US/09009913  
Patent No. 6087485  
GENERAL INFORMATION:  
APPLICANT: Axxs Pharmaceuticals, Inc.  
TITLE OF INVENTION: Asthma Related Genes  
NUMBER OF SEQUENCES: 339  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bozicevic & Reed, LLP  
STREET: 285 Hamilton Ave, Suite 200  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94301

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/009,913  
FILING DATE: 21-JAN-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Sherwood, Pamela J  
REGISTRATION NUMBER: 36,677  
REFERENCE/DOCKET NUMBER: SEQ-4P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-327-3231  
TELEFAX: 650-327-3231  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 72928 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
US-09-009-913-1  
Alignment Scores:  
Pred. No.: 94.9 Length: 72928  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.28% Indels: 0  
Gaps: 0  
DB: 3  
US-09-854-133-586 (1-97) x US-09-009-913-1 (1-72928)  
QY 1 GluValGluValSerArgAspHisAla 9  
|||||  
Db 47805 GAGTTGAAGTGAGCAGAGATCACGCC 47779  
RESULT 11  
US-09-128-155-16  
Sequence 16, Application US/09128155  
Patent No. 6117654  
GENERAL INFORMATION:  
APPLICANT: Pan, Yang  
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY  
TITLE OF INVENTION: AND USES THEREOF  
FILE REFERENCE: 09404/052001  
CURRENT APPLICATION NUMBER: US/09/128,155  
CURRENT FILING DATE: 1998-08-03  
EARLIER APPLICATION NUMBER: US 60/091,650  
EARLIER FILING DATE: 1998-07-02  
EARLIER APPLICATION NUMBER: US 60/054,646  
EARLIER FILING DATE: 1997-08-04  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 16  
LENGTH: 152331  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(152331)  
OTHER INFORMATION: n = A,T,C or G  
US-09-128-155-16  
Alignment Scores:  
Pred. No.: 188 Length: 152331  
Score: 9.00 Matches: 9

Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 9.28%  
DB: 3  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-09-128-155-16 (1-152331)

QY 10 SerLeuGLYAspSerGLuThrLeuSer 18

Db 118192 AGCCTGGGTGACAGCAGACCTTGTCT 118218

## RESULT 12

US-08-004-800-16/c

Sequence 16, Application US/08004800

Patent No. 5426180

GENERAL INFORMATION:

APPLICANT: KOOL, Eric T.

TITLE OF INVENTION: SINGLE-STRANDED, CIRCULAR

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Scully, Scott, Murphy & Presser

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: New York

COUNTRY: USA

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/004,800

FILING DATE: 19930111

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: McNulty, William E.

REGISTRATION NUMBER: 22,606

REFERENCE/DOCKET NUMBER: 80852ZY

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343

TELEFAX: (516) 742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 36 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

US-08-004-800-16

## Alignment Scores:

Pred. No.: 0.84

Score: 8.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 8.25%

Length: 36  
Matches: 8  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-08-004-800-16 (1-36)

QY 27 ArgLysLysLysArgGLuArgLys 34

Db 35 AGAAAAAGAGAGAGAGAGAAAA 12

## RESULT 13

US-08-413-813-16/c

Sequence 16, Application US/08413813

Patent No. 5683874

GENERAL INFORMATION:

APPLICANT: KOOL, Eric T.

TITLE OF INVENTION: SINGLE-STRANDED, CIRCULAR OLIGONUCLEOTIDES  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Scully, Scott, Murphy & Presser

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: New York

COUNTRY: USA

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/413,813

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Digilio, Frank S.

REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 80852ZYX

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343

TELEFAX: (516) 742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 36 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-413-813-16

## Alignment Scores:

Pred. No.: 0.84

Score: 8.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 8.25%

Length: 36  
Matches: 8  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-08-413-813-16 (1-36)

QY 27 ArgLysLysLysArgGLuArgLys 34

Db 35 AGAAAAAGAGAGAGAGAGAAAA 12

## RESULT 14

US-08-467-346-16/c

Sequence 16, Application US/08467346

Patent No. 5872105

GENERAL INFORMATION:

APPLICANT: KOOL, Eric T.

TITLE OF INVENTION: SINGLE-STRANDED, CIRCULAR OLIGONUCLEOTIDES

NUMBER OF SEQUENCES: 44

CORRESPONDENCE ADDRESS:

ADDRESSEE: Scully, Scott, Murphy & Presser

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: New York

COUNTRY: USA

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,346

FILING DATE: 06-JUN-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/413,813  
FILING DATE: 30-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Digilio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 8085ZYX  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-467-346-16

Alignment Scores:  
Pred. No.: 0.84 Length: 36  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 2 Gaps: 0

US-09-854-133-586 (1-97) x US-08-467-346-16 (1-36)

OY 27 ArglyslslysgluArglyslslys 34  
Db 35 AGAAAAAGAAAGAGAGAGAGAAAA 12

## RESULT 15

US-09-084-120-18/c  
Sequence 18, Application US/09084120  
Patent No. 6251592  
GENERAL INFORMATION:  
APPLICANT: TANG, JianQing  
APPLICANT: MELANCON, Serge B.  
TITLE OF INVENTION: A NOVEL STR MARKER SYSTEM  
TITLE OF INVENTION: FOR DNA FINGERPRINTING  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SWABEY OGILVY RENAULT  
STREET: Suite 1600, 1981 McGill College Avenue  
CITY: Montreal  
STATE: QC  
COUNTRY: Canada  
ZIP: H3A 2Y3  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FastSeq for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/084,120  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: COTE, France  
REGISTRATION NUMBER: 37,037  
REFERENCE/DOCKET NUMBER: 13251-4US FC/1d  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 514 845-7126  
TELEFAX: 514 288-8389  
TELEX:  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 139 base pairs  
TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
US-09-084-120-18

Alignment Scores:  
Pred. No.: 2.95 Length: 139  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 4 Gaps: 0

US-09-854-133-586 (1-97) x US-09-084-120-18 (1-139)

OY 23 ArglyslslysgluArglyslslys 30  
Db 122 AGAAAGAAAGAAAGAAAGAAAGAA 99

## RESULT 16

US-07-922-723A-7  
Sequence 7, Application US/07922723A  
Patent No. 5369004  
GENERAL INFORMATION:  
APPLICANT: Drs. Mihael H. Polymeropoulos  
APPLICANT: and Carl R. Merrill  
TITLE OF INVENTION: FIVE HIGHLY INFORMATIVE  
TITLE OF INVENTION: REPEAT POLYMORPHIC DNA MARKERS  
NUMBER OF SEQUENCES: 73  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lowe, Price, LeBlanc & Becker  
STREET: Suite 300, 99 Canal Center Plaza  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22314  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: DOS Text File  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/922,723A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: D.J. Mills  
REGISTRATION NUMBER: 34506  
REFERENCE/DOCKET NUMBER: 717081B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 684 1111  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 291  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-07-922-723A-7

Alignment Scores:  
Pred. No.: 5.87 Length: 291  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 1 Gaps: 0

US-09-854-133-586 (1-97) x US-07-922-723A-7 (1-291)

OY 23 ArglyslslysgluArglyslslys 30  
Db 169 AGAAAGAAAGAAAGAAAGAAAG 192





TELEPHONE: 703 684 1111  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 291  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-480-366-7

Alignment Scores:  
Pred. No.: 5.87 Length: 291  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 1 Gaps: 0

US-09-854-133-586 (1-97) x US-08-480-366-7 (1-291)

OY 23 ArgLysLysGluArgLysLysLys 30  
Db 169 AGAAGAGAAAGAGAGAGAGAGAGAG 192

RESULT 20  
US-07-952-277A-7

; Sequence 7, Application US/07952277A

; Patent No. 5861504

; GENERAL INFORMATION:

APPLICANT: Drs. Michael H. Polymeropoulos

APPLICANT: and Carl R. Merrill

TITLE OF INVENTION: ELEVEN HIGHLY INFORMATIVE

TITLE OF INVENTION: REPEAT POLYMORPHIC DNA MARKERS

NUMBER OF SEQUENCES: 85

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lowe, Price, LeBlanc & Becker

STREET: Suite 300, 99 Canal Center Plaza

CITY: Alexandria

STATE: Virginia

COUNTRY: USA

ZIP: 22314

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: DOS Text File

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/952,277A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: D.J. Mills

REGISTRATION NUMBER: 34506

REFERENCE/DOCKET NUMBER: 717081C

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703 684 1111

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 291

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-07-952-277A-7

Alignment Scores:

Pred. No.: 5.87 Length: 291

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 8.25% Indels: 0

DB: 2 Gaps: 0

US-09-854-133-586 (1-97) x US-07-952-277A-7 (1-291)

OY 23 ArgLysLysGluArgLysLysLys 30  
Db 169 AGAAGAGAAAGAGAGAGAGAGAGAG 192

RESULT 21

US-08-623-906A-19

; Sequence 19, Application US/08623906A

; Patent No. 5874217

; GENERAL INFORMATION:

APPLICANT: Stevenson, Tamara

APPLICANT: Dvorak, Jan

APPLICANT: Halverson, Joy

TITLE OF INVENTION: Microsatellite Sequences for Canine

TITLE OF INVENTION: Genotyping

NUMBER OF SEQUENCES: 60

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: CA

COUNTRY: US

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/623,906A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Sherwood, Pamela J.

REGISTRATION NUMBER: 36,677

REFERENCE/DOCKET NUMBER: A-62282/BIR

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-781-1989

TELEFAX: 415-398-3249

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 299 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: misc\_feature

LOCATION: 1..128

OTHER INFORMATION: /note= "Nucleotides 1-128 are

OTHER INFORMATION: unique flanking sequence"

FEATURE:

NAME/KEY: misc\_feature

LOCATION: 129..199

OTHER INFORMATION: /note= "Nucleotides 129-199 are

OTHER INFORMATION: repeat sequence"

FEATURE:

NAME/KEY: misc\_feature

LOCATION: 200..299

OTHER INFORMATION: /note= "Nucleotides 200-299 are

OTHER INFORMATION: unique flanking sequence"

US-08-623-906A-19

Alignment Scores:

Pred. No.: 6.02 Length: 299

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 8.25% Indels: 0

DB: 2 Gaps: 0

US-09-854-133-586 (1-97) x US-08-623-906A-19 (1-299)

OY 23 ArglyslgsluArglysls 30  
Db 176 AGAAGAAAGAAAGAAAGAAAG 199

## RESULT 22

US-08-623-906A-17

; Sequence 17, Application US/08623906A

; Patent No. 5874217

; GENERAL INFORMATION:

; APPLICANT: Stevenson, Tamara

; APPLICANT: Dvorak, Jan

; APPLICANT: Halverson, Joy

; TITLE OF INVENTION: Microsatellite Sequences for Canine

; NUMBER OF SEQUENCES: 60

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON &amp; HERBERT

; STREET: 4 Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: CA

; COUNTRY: US

; ZIP: 94111-4187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/623,906A

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Sherwood, Pamela J.

; REGISTRATION NUMBER: 36,677

; REFERENCE/DOCKET NUMBER: A-62282/BIR

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-781-1989

; TELEFAX: 415-398-3249

; INFORMATION FOR SEQ ID NO: 17:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 350 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: 1..109

; OTHER INFORMATION: /note= "Nucleotides 1-109 are

; OTHER INFORMATION: unique flanking sequence"

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: 110..195

; OTHER INFORMATION: /note= "Nucleotides 110-195 are

; OTHER INFORMATION: repeat sequence"

; NAME/KEY: misc\_feature

; LOCATION: 196..350

; OTHER INFORMATION: /note= "Nucleotides 196-350 are

; OTHER INFORMATION: unique flanking sequence"

US-08-623-906A-17

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-854-133-586 (1-97) x US-08-623-906A-17 (1-350)

OY 23 ArglyslgsluArglysls 30  
Db 118 AGAAGAAAGAAAGAAAGAAAG 141

## RESULT 23

US-09-157-177-110

; Sequence 110, Application US/09157177

; Patent No. 6090558

; GENERAL INFORMATION:

; APPLICANT: Butler, John M.

; APPLICANT: Li, Jia

; APPLICANT: Monforte, Joseph A.

; APPLICANT: Becker, Christopher H.

; TITLE OF INVENTION: DNA TYPING BY MASS SPECTROMETRY WITH POLYMORPHIC DNA

; FILE REFERENCE: GETR:017/GETR017P

; CURRENT APPLICATION NUMBER: US/09/157,177

; EARLIER FILING DATE: 1998-09-18

; EARLIER FILING DATE: 1997-09-19

; NUMBER OF SEQ ID NOS: 135

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 110

; LENGTH: 350

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-157-177-110

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-854-133-586 (1-97) x US-09-157-177-110 (1-350)

OY 22 LeuArglyslgsluArglysls 29  
Db 70 CTCAGAAAGAAAGAAAGAAAG 93

## RESULT 24

US-08-623-906A-18

; Sequence 18, Application US/08623906A

; Patent No. 5874217

; GENERAL INFORMATION:

; APPLICANT: Stevenson, Tamara

; APPLICANT: Dvorak, Jan

; APPLICANT: Halverson, Joy

; TITLE OF INVENTION: Microsatellite Sequences for Canine

; NUMBER OF SEQUENCES: 60

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON &amp; HERBERT

; STREET: 4 Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: CA

; COUNTRY: US

; ZIP: 94111-4187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/623,906A

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Sherwood, Pamela J.

; REGISTRATION NUMBER: 36,677

; REFERENCE/DOCKET NUMBER: A-62282/BIR

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 376 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..79  
OTHER INFORMATION: /note= "Nucleotides 1-79 are unique  
flanking sequence"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 80..229  
OTHER INFORMATION: /note= "Nucleotides 80-229 are  
repeat sequence"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 230..376  
OTHER INFORMATION: /note= "Nucleotides 230-376 are  
unique flanking sequence"  
US-08-623-906A-18

Alignment Scores:  
Pred. No.: 7.45 Length: 376  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 2 Gaps: 0

US-09-854-133-586 (1-97) x US-08-623-906A-18 (1-376)

QY 23 Arglyslsgluarglyslslys 30  
|||||  
Db 98 AGAAGAGAAGAGAAGAGAAGAAA 121

RESULT 25  
US-08-623-906A-13/c  
Sequence 13, Application US/08623906A  
Patent No. 5874217  
GENERAL INFORMATION:  
APPLICANT: Stevenson, Tamara  
APPLICANT: Dvorak, Jan  
APPLICANT: Halverson, Joy  
TITLE OF INVENTION: Microsatellite Sequences for Canine  
TITLE OF INVENTION: Genotyping  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/623,906A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Sherwood, Pamela J.  
REGISTRATION NUMBER: 36,677  
REFERENCE/DOCKET NUMBER: A-62282/BIR  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 388 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..131  
OTHER INFORMATION: /note= "Nucleotides 1-131 are  
unique flanking sequence"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 132..235  
OTHER INFORMATION: /note= "Nucleotides 132-235 are  
repeat sequence"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 236..388  
OTHER INFORMATION: /note= "Nucleotides 236-388 are  
unique flanking sequence"  
US-08-623-906A-13

Alignment Scores:  
Pred. No.: 7.67 Length: 388  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 2 Gaps: 0

US-09-854-133-586 (1-97) x US-08-623-906A-13 (1-388)

QY 23 Arglyslsgluarglyslslys 30  
|||||  
Db 176 AGAAGAGAAGAGAAGAGAAGAAA 153

RESULT 26  
US-09-018-584A-22  
Sequence 22, Application US/09018584A  
Patent No. 6238863  
GENERAL INFORMATION:  
APPLICANT: Schumm, James W.  
APPLICANT: Bacher, Jeffery W.  
TITLE OF INVENTION: MATERIALS AND METHODS FOR  
IDENTIFYING AND ANALYZING INTERMEDIATE TANDEM  
REPEAT DNA MARKERS  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Promega Corporation  
STREET: 2800 Woods Hollow Road  
CITY: Madison  
STATE: Wisconsin  
COUNTRY: U.S.A.  
ZIP: 53711-5399  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette - 3.5 inch, 1.44 MB  
COMPUTER: IBM compatible PC  
OPERATING SYSTEM: Windows 95  
SOFTWARE: Word 97 (DOS text format)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/018,584A  
FILING DATE: 04-Feb-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Grady J. Frenchick  
REGISTRATION NUMBER: 29,018  
REFERENCE/DOCKET NUMBER: 16026.9180  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (608) 257-3501

TELEFAX: (608) 257-2275  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 412 bp  
; TYPE: Nucleic Acid  
; STRANDEDNESS: Double  
; TOPOLOGY: Circular  
; MOLECULE TYPE: Genomic DNA  
; HYPOTHETICAL: no  
; IMMEDIATE SOURCE:  
; LIBRARY: plasmid, pGem3zf(+)  
; CLONE: G234  
; POSITION IN GENOME:  
; CHROMOSOME/SEGMENT: 16 qter  
US-09-018-584A-22

Alignment Scores:  
Pred. No.: 8.11 Length: 412  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 4 Gaps: 0

US-09-854-133-586 (1-97) x US-09-018-584A-22 (1-412)

OY 22 leuArglyslsGluArglysls 29  
Db 49 CTCAGAAAAAGAAAGAAAGAAAGAA 72

## RESULT 27

US-08-623-906A-6  
; Sequence 6, Application US/08623906A  
; Patent No. 5874217

## GENERAL INFORMATION:

APPLICANT: Stevenson, Tamara  
APPLICANT: Dvorak, Jan  
APPLICANT: Halverson, Joy  
TITLE OF INVENTION: Microsatellite Sequences for Canine  
TITLE OF INVENTION: Genotyping  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA

COUNTRY: US

ZIP: 94111-4187

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/623,906A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Sherwood, Pamela J.

REGISTRATION NUMBER: 36,677

REFERENCE/DOCKET NUMBER: A-62282/BIR

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-781-1989

TELEFAX: 415-398-3249

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 454 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:  
NAME/KEY: misc\_feature

LOCATION: 1..133  
; OTHER INFORMATION: /note= "Nucleotides 1-133 are  
; OTHER INFORMATION: unique flanking sequence"  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 134..307  
; OTHER INFORMATION: /note= "Nucleotides 134-207 are  
; OTHER INFORMATION: repeat sequence"  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 308..454  
; OTHER INFORMATION: /note= "Nucleotides 308-454 are  
; OTHER INFORMATION: unique flanking sequence"  
US-08-623-906A-6

Alignment Scores:  
Pred. No.: 8.87 Length: 454  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 2 Gaps: 0

US-09-854-133-586 (1-97) x US-08-623-906A-6 (1-454)

OY 23 ArglyslsGluArglysls 30  
Db 216 AGAAAGAAAGAAAGAAAGAAAG 239

## RESULT 28

US-09-385-982-31  
; Sequence 31, Application US/09385982  
; Patent No. 6262334

## GENERAL INFORMATION:

APPLICANT: ENDEGE, WILSON O., ET AL.  
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION  
TITLE OF INVENTION: PRODUCTS: II  
FILE REFERENCE: CCDNA-260XX  
CURRENT APPLICATION NUMBER: US/09/385,982  
CURRENT FILING DATE: 1999-08-30  
EARLIER APPLICATION NUMBER: 09/328,111  
EARLIER FILING DATE: 1999-06-08  
EARLIER APPLICATION NUMBER: 60/117,393  
EARLIER FILING DATE: 1999-01-27  
EARLIER APPLICATION NUMBER: 60/098,639  
EARLIER FILING DATE: 1998-08-31  
NUMBER OF SEQ ID NOS: 544  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 31  
LENGTH: 485  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)..(485)  
OTHER INFORMATION: n = A,T,C or G  
US-09-385-982-31

Alignment Scores:  
Pred. No.: 9.43 Length: 485  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 4 Gaps: 0

US-09-854-133-586 (1-97) x US-09-385-982-31 (1-485)

OY 23 ArglyslsGluArglysls 30  
Db 136 AGAAAGAAAGAAAGAAAGAAAG 159

## RESULT 29



```

US-08-332-766A-22
; Sequence 22, Application US/08332766A
; Patent No. 5843647
;
; GENERAL INFORMATION:
; APPLICANT: JEFFREYS, Alec J.
; APPLICANT: ARMOUR, John
; TITLE OF INVENTION: SIMPLE TANDEM REPEATS
; NUMBER OF SEQUENCES: 125
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN, L.L.P.
; STREET: 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: U.S.A.
; ZIP: 20005-3918
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,766A
; FILING DATE: 01-NOV-1994
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9326052.9
; FILING DATE: 21-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BIRD, Donald J.
; REGISTRATION NUMBER: 25,323
; REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
;
US-08-332-766A-22
;
Alignment Scores:
Pred. No.: 9.6 Length: 494
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.25% Indels: 0
DB: 2 Gaps: 0
;
US-09-854-133-586 (1-97) x US-08-332-766A-22 (1-494)
OY 23 ArgLysLysGluArgLysLys 30
|||||
Db 108 AGAAGAAAGAAAGAAAGAA 131
|||||
RESULT 30
US-09-103-359-4
; Sequence 4, Application US/09103359
; Patent No. 6057108
;
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; APPLICANT: Patterson, Chandra
; TITLE OF INVENTION: HUMAN ARF-RELATED PROTEINS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.

```

```

: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows
: SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/103,359
: FILING DATE: HEREWITH
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Cerrone, Michael C
: REGISTRATION NUMBER: 39,132
: REFERENCE/DOCKET NUMBER: PF-0537 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-855-0555
: TELEFAX: 650-855-0572
:
: TELEX:
:
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1005 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: COLNNOT13
: CLONE: 1333754
:
: US-09-103-359-4
:
: Alignment Scores:
: Pred. No.: 18.6 Length: 1005
: Score: 8.00 Matches: 8
: Percent Similarity: 100.00% Conservative: 0
: Best Local Similarity: 100.00% Mismatches: 0
: Query Match: 8.25% Indels: 0
: DB: 3 Gaps: 0
:
: US-09-854-133-586 (1-97) x US-09-103-359-4 (1-1005)
:
: QY 23 ArgLysLysGluArgLysLysLys 30
: |||||
: Db 115 AGAAGAAAGAAAGAAAGAAAAA 138
:
: RESULT 31
: US-08-097-938-3
: Sequence 3, Application US/08097938
: Patent No. 5629174
: GENERAL INFORMATION:
: APPLICANT: SUNDELIN, JOHAN
: APPLICANT: SCARBOROUGH, ROBERT M.
: TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
: TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
: NUMBER OF SEQUENCES: 59
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: MORRISON & FOERSTER
: STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
: CITY: Washington, D.C.
: COUNTRY: USA
: ZIP: 20006-1812
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/097,938

```

FILING DATE: 26-JUL-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 22803-20006.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..1249  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 56  
US-08-097-938-3

Alignment Scores:  
Pred. No.: 22.8 Length: 1255  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 1 Gaps: 0

US-09-854-133-586 (1-97) x US-08-097-938-3 (1-1255)

QY 11 LeuGlyAspSerGluThrLeuSer 18  
Db 11 CTGGTGACAGCAGACCCCTGTCT 34

## RESULT 32

US-08-476-000-3  
Sequence 3, Application US/08476000  
Patent No. 5716789

GENERAL INFORMATION:  
APPLICANT: SUNDELIN, JOHAN  
APPLICANT: SCARBOROUGH, ROBERT M.  
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS  
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20006-1812

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,000  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/390,301  
FILING DATE: 25-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ADLER, REID G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 2803-0006.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..1249  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 56  
US-08-476-000-3

Alignment Scores:  
Pred. No.: 22.8 Length: 1255  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 1 Gaps: 0

US-09-854-133-586 (1-97) x US-08-476-000-3 (1-1255)

QY 11 LeuGlyAspSerGluThrLeuSer 18  
Db 11 CTGGTGACAGCAGACCCCTGTCT 34

## RESULT 33

US-08-472-840-3  
Sequence 3, Application US/08472840  
Patent No. 5763575

GENERAL INFORMATION:  
APPLICANT: SUNDELIN, JOHAN  
APPLICANT: SCARBOROUGH, ROBERT M.  
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS  
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20006-1812

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,840  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/390,301  
FILING DATE: 25-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ADLER, REID G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 2803-0006.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..1249  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 56  
US-08-472-840-3

Alignment Scores:  
Pred. No.: 22.8  
Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 8.25%  
DB: 1  
Length: 1255  
Matches: 8  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-08-472-840-3 (1-1255)

QY 11 LeuGlyAspSerGluThrLeuSer 18  
|||||  
Db 11 CTGGGTGACAGCGAGACCTGTCT 34

RESULT 34  
US-08-476-976-3  
Sequence 3, Application US/08476976  
Patent No. 5874400

GENERAL INFORMATION:  
APPLICANT: SUNDELIN, JOHAN  
APPLICANT: SCARBOROUGH, ROBERT M.  
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS  
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20006-1812

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,976  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/390,301  
FILING DATE: 25-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ADLER, REID G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 2803-0006.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..1249  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 56  
US-08-476-976-3

Alignment Scores:  
Pred. No.: 22.8  
Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 8.25%  
DB: 2  
Length: 1255  
Matches: 8  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-08-476-976-3 (1-1255)

QY 11 LeuGlyAspSerGluThrLeuSer 18  
|||||  
Db 11 CTGGGTGACAGCGAGACCTGTCT 34

RESULT 35  
US-08-474-410-3  
Sequence 3, Application US/08474410  
Patent No. 6043212

GENERAL INFORMATION:  
APPLICANT: SUNDELIN, JOHAN  
APPLICANT: SCARBOROUGH, ROBERT M.  
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS  
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20006-1812

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,410  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/390,301  
FILING DATE: 25-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ADLER, REID G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 2803-0006.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..1249  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 56  
US-08-474-410-3

Alignment Scores:  
Pred. No.: 22.8  
Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 8.25%  
DB: 3  
Length: 1255  
Matches: 8  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-08-474-410-3 (1-1255)

QY 11 LeuGlyaspSergluThrleuser 18  
DB 11 CTGGGTGACGACGAGACCTGTCT 34

## RESULT 36

US-08-486-673B-3

; Sequence 3, Application US/08486673B  
; Patent No. 6297026

; GENERAL INFORMATION:

; APPLICANT: Sundelin, Johan

; APPLICANT: Scarborough, Robert M.

; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor

; FILE REFERENCE: 44481-5006-08-US

; CURRENT APPLICATION NUMBER: US/08/486,673B

; PRIOR FILING DATE: 1995-06-07

; PRIOR APPLICATION NUMBER: US 08/097,938

; PRIOR FILING DATE: 1994-07-26

; NUMBER OF SEQ ID NOS: 63

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 1255

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (56)..(1249)

; OTHER INFORMATION: C140 receptor, genomic DNA and deduced protein

; OTHER INFORMATION: sequences

US-08-486-673B-3

Alignment Scores:  
Pred. No.: 22.8  
Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 8.25%  
DB: 4  
Length: 1255  
Matches: 8  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-08-486-673B-3 (1-1255)

QY 11 LeuGlyaspSergluThrleuser 18  
DB 11 CTGGGTGACGACGAGACCTGTCT 34

## RESULT 37

US-09-184-001-1/C

; Sequence 1, Application US/09184001A  
; Patent No. 6303333

; GENERAL INFORMATION:

; APPLICANT: MICHALOVICH, DAVID

; APPLICANT: SIMS, MATTHEW A.

; APPLICANT: SHAIKH, NARJIS

; TITLE OF INVENTION: NOVEL COMPOUNDS

; FILE REFERENCE: GP-30088

; CURRENT APPLICATION NUMBER: US/09/184,001A

; EARLIER FILING DATE: 1998-11-02

; EARLIER APPLICATION NUMBER: 9817479.0

; EARLIER FILING DATE: 1998-08-11

; EARLIER APPLICATION NUMBER: 9806221.9

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1

; LENGTH: 2186

; TYPE: DNA

; ORGANISM: HOMO SAPIENS

US-09-184-001-1

Pred. No.: 38.3  
Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 8.25%  
DB: 4  
Length: 2186  
Matches: 8  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-09-184-001-1 (1-2186)

QY 87 ThirArgSerHisLeuGlyArg 94  
DB 215 ACCAGAGACCATCTGGGAGAGAGG 192

## RESULT 38

US-08-458-434A-1

; Sequence 1, Application US/08458434A  
; Patent No. 6083690

; GENERAL INFORMATION:

; APPLICANT: Harris Ph.D., Stephen E.

; APPLICANT: Mundy M.D., Gregory R.

; APPLICANT: Gosh-Choudhury Ph.D., Nandini

; APPLICANT: Feng Ph.D., Jian O.

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IDENTIFYING

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: James C. Weseman, Esq.

; STREET: 401 B. Street, Suite 1700

; CITY: San Diego

; STATE: CA

; COUNTRY: USA

; ZIP: 92101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/458,434A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Weseman, James C.

REGISTRATION NUMBER: 30,507

REFERENCE/DOCKET NUMBER: P00060US0

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 699-3604

TELEFAX: 619-236-1048

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2310 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 768..1991

US-08-458-434A-1

US-09-854-133-586 (1-97) x US-08-458-434A-1 (1-2310)

QY 23 ArgLysLysGluArgLysLys 30  
DB 2143 AGAAGAAAGAAAGAAAGAAAGAA 2166

Alignment Scores:  
Pred. No.: 40.3  
Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 8.25%  
DB: 3  
Length: 2310  
Matches: 8  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0



RESULT 39  
US-09-019-689-1/c  
; Sequence 1, Application US/09019689  
; Patent No. 6103527  
; GENERAL INFORMATION:  
; APPLICANT: GRAULICH, Wolff  
; APPLICANT: NETTELBECK, Dirk  
; APPLICANT: SEDLACEK, Hans-Harald  
; APPLICANT: MUELLER, Rolf  
; TITLE OF INVENTION: PROMOTER OF THE HUMAN ENDOGLIN GENE AND  
; TITLE OF INVENTION: ITS USE  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY & LARDNER  
; STREET: 3000 K Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/019,689  
; FILING DATE: 06-FEB-1998  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 19704301.1  
; FILING DATE: 06-FEB-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sandercock, Colin G.  
; REGISTRATION NUMBER: 31,298  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 672-5300  
; TELEFAX: (202) 672-5399  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2415 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-09-019-689-1  
  
Alignment Scores:  
Pred. No.: 42 Length: 2415  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 3 Gaps: 0  
  
US-09-854-133-586 (1-97) x US-09-019-689-1 (1-2415)  
  
QY 24 LysLysGluArgLysLysLysArg 31  
DB 1381 AAAAAAGAAAGAAAAAAGAGA 1358  
  
RESULT 40  
US-09-184-001-3/c  
; Sequence 3, Application US/09184001A  
; Patent No. 6303333  
; GENERAL INFORMATION:  
; APPLICANT: MICHALOVICH, DAVID  
; APPLICANT: SIMS, MATTHEW A.  
; APPLICANT: SHAIKH, NARJIS  
; TITLE OF INVENTION: NOVEL COMPOUNDS  
; FILE REFERENCE: GP-30088  
; CURRENT APPLICATION NUMBER: US/09/184,001A

; CURRENT FILING DATE: 1998-11-02  
; EARLIER APPLICATION NUMBER: 9817479.0  
; EARLIER FILING DATE: 1998-08-11  
; EARLIER APPLICATION NUMBER: 9806221.9  
; EARLIER FILING DATE: 1998-03-23  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 2558  
; TYPE: DNA  
; ORGANISM: HOMO SAPIENS  
; FEATURE:  
; NAME/KEY: UNSURE  
; LOCATION: (146)(161)(178)(233)(254)(296)  
; US-09-184-001-3  
  
Alignment Scores:  
Pred. No.: 44.3 Length: 2558  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 4 Gaps: 0  
  
US-09-854-133-586 (1-97) x US-09-184-001-3 (1-2558)  
  
QY 87 ThrArgSerHisLeuGlyArgArg 94  
DB 111 ACCAGAGCACACTGGGAGAGAAG 88  
  
RESULT 41  
US-08-555-723B-3/c  
; Sequence 3, Application US/08555723B  
; Patent No. 5837534  
; GENERAL INFORMATION:  
; APPLICANT: OLSON, Eric N.  
; APPLICANT: LI, Li  
; APPLICANT: MIANO, Joseph M.  
; TITLE OF INVENTION: SMOOTH MUSCLE 22' PROMOTER, GENE TRANSFER  
; TITLE OF INVENTION: VECTORS CONTAINING THE SAME, AND METHOD OF  
; TITLE OF INVENTION: USE OF THE SAME TO TARGET GENE EXPRESSION IN  
; TITLE OF INVENTION: ARTERIAL SMOOTH MUSCLE CELLS  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
; STREET: 2100 Pennsylvania Avenue, N.W., Suite 800  
; CITY: Washington, D.C.  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/555,723B  
; FILING DATE: 14-NOVEMBER-1995  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KIT, Gordon  
; REGISTRATION NUMBER: 30,764  
; REFERENCE/DOCKET NUMBER: A-6663  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 293-7060  
; TELEFAX: (202) 293-7860  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3892 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: genomic

; HYPOTHETICAL: NO  
US-08-555-723B-3

Alignment Scores:  
Pred. No.: 65.5  
Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 8.25%  
DB: 2

Length: 3892  
Matches: 8  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-08-555-723B-3 (1-3892)

QY 27 ArgLysLysLysArgGluArgLys 34  
DB 3448 AGAAGAGAAAAAGAGAGAGAAAA 3425

RESULT 42  
US-09-123-465-3/C  
; Sequence 3, Application US/09123465  
; Patent No. 6015711  
; GENERAL INFORMATION:  
; APPLICANT: OLSON, Eric N.  
; APPLICANT: LI, Li  
; APPLICANT: MIANO, Joseph M.  
; TITLE OF INVENTION: SMOOTH MUSCLE 22' PROMOTER, GENE TRANSFER  
; TITLE OF INVENTION: VECTORS CONTAINING THE SAME, AND METHOD OF  
; TITLE OF INVENTION: USE OF THE SAME TO TARGET GENE EXPRESSION IN  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
; STREET: 2100 Pennsylvania Avenue, N.W., Suite 800  
; CITY: Washington, D.C.  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20037

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/123,465  
; FILING DATE: 28-JULY-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/555,723  
; FILING DATE: 14-NOVEMBER-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KIT, Gordon  
; REGISTRATION NUMBER: 30,764  
; REFERENCE/DOCKET NUMBER: A-7243  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 293-7060  
; TELEFAX: (202) 293-7860  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3892 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: genomic  
; HYPOTHETICAL: NO  
; US-09-123-465-3

Alignment Scores:  
Pred. No.: 65.5  
Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 8.25%

Length: 3892  
Matches: 8  
Conservative: 0  
Mismatch: 0  
Indels: 0

DB: 3 Gaps: 0

US-09-854-133-586 (1-97) x US-09-123-465-3 (1-3892)

QY 27 ArgLysLysLysArgGluArgLys 34  
DB 3448 AGAAGAGAAAAAGAGAGAGAAAA 3425

RESULT 43  
US-09-122-126B-1/C  
; Sequence 1, Application US/09122126B  
; Patent No. 6451575  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: AGGREGAN DEGRADING METALLO PROTEASES  
; FILE REFERENCE: DM6909  
; CURRENT APPLICATION NUMBER: US/09/122,126B  
; NUMBER OF SEQ ID NOS: 21  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 1  
; LENGTH: 4192  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (406)..(2916)  
US-09-122-126B-1

Alignment Scores:  
Pred. No.: 70.1  
Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 8.25%  
DB: 4

Length: 4192  
Matches: 8  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-09-122-126B-1 (1-4192)

QY 23 ArgLysLysGluArgLysLys 30  
DB 3533 AGAAGAGAAAAAGAGAGAAAAAA 3510

RESULT 44  
US-08-458-434A-7  
; Sequence 7, Application US/08458434A  
; Patent No. 6083690  
; GENERAL INFORMATION:  
; APPLICANT: Harris Ph.D., Stephen E.  
; APPLICANT: Mundy M.D., Gregory R.  
; APPLICANT: Gosh-Choudhury Ph.D., Nandini  
; APPLICANT: Feng Ph.D., Jian Q.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IDENTIFYING  
; NUMBER OF SEQUENCES: 13  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: James C. Weseman, Esq.  
; STREET: 401 B. Street, Suite 1700  
; CITY: San Diego  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92101

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/458,434A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weseman, James C.

REGISTRATION NUMBER: 30,507  
REFERENCE/DOCKET NUMBER: P00060US0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 699-3604  
TELEFAX: 619-236-1048  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9299 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-458-434A-7

Alignment Scores:  
Pred. No.: 147  
Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 8.25%  
DB: 3

Length: 9299  
Matches: 8  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-08-458-434A-7 (1-9299)

QY 23 ArgLysLysGluArgLysLysLys 30  
Db 9134 AGAAGAAAGAAAGAAAGAAAGAA 9157

RESULT 45  
US-08-669-161A-29/C  
Sequence 29, Application US/08669161A  
Patent No. 6013481  
GENERAL INFORMATION:  
APPLICANT: Debacker, Olivier; Van den Eynde,  
APPLICANT: Benoit; Boon-Falleur, Thierry  
TITLE OF INVENTION: Isolated, Truncated Nucleic Acid  
TITLE OF INVENTION: Molecules Which Code For GAGE Tumor Rejection Antigen,  
TITLE OF INVENTION: The Tumor Rejection Antigen, and Uses Thereof  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/669,161A  
FILING DATE: 24-June-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/531,662  
FILING DATE: 21-September-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/370,648  
FILING DATE: 10-January-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/250,162  
FILING DATE: 27-May-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/096,039  
FILING DATE: 22-July-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 6013481man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: LUD 5443  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11461 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-669-161A-29

Alignment Scores:  
Pred. No.: 179  
Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 8.25%  
DB: 3

Length: 11461  
Matches: 8  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-08-669-161A-29 (1-11461)

QY 24 LysLysGluArgLysLysLysArg 31  
Db 10461 AAGAAAGAAAGAAAGAAAGAAAG 10438

Search completed: May 11, 2003, 17:17:42  
Job time : 87.0797 secs







;; CURRENT APPLICATION NUMBER: US/09/854,133  
;; CURRENT FILING DATE: 2001-05-11  
;; NUMBER OF SEQ ID NOS: 735  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 442  
;; LENGTH: 337  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-854-133-442

## Alignment Scores:

Pred. No.: 3.54e-94 Length: 337  
Score: 97.00 Matches: 97  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-09-854-133-586 (1-97) x US-09-854-133-442 (1-337)

QY 1 GluValGluValSerArgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThr 20  
Db 5 GAGTTGAAGTGAGCAGAGATCATGCCCTGGGTGACAGTGAAGACTCTGTCTCAACA 64  
QY 21 GluLeuArgLysLysGluArgLysLysArgGluArgLysPheGlnAlaAsnCysGly 40  
Db 65 GAATTAAGGAAAAAGAAAGAAAAAGAGAGAGAGAAATCCAGGCCAATTGTGGC 124  
QY 41 IleAspPheIleIlePheTrpIlePheTrpIleLeuLeuPheSerHisTrpIleGln 60  
Db 125 ATAGATTATCATATCTGTGATTTTGGATTCTTTGTTCTCATCACTGATTTCAG 184  
QY 61 GluSerLeuLeuCysProProSerProLysGluValThrCysArgGluMetLeuThrGly 80  
Db 185 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGAGTTACTGCAGGGAATGTAAACGGGA 244  
QY 81 GlyCysLeuProTrpAlaThrArgSerHisLeuGlyArgArgLysCysSer 97  
Db 245 GGCTGCCTTCCCTGGGCAACAAGAGCCACTGGGCGAGAGAAAGTGACGC 295

## RESULT 2

US-09-738-973-442

;; Sequence 442, Application US/09738973  
;; Patent No. US20020110563A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Reed, Steven G.  
;; APPLICANT: Henderson, Robert A.  
;; APPLICANT: Lodes, Michael J.  
;; APPLICANT: Flinn, Steven P.  
;; APPLICANT: Mohamath, Raodoh  
;; APPLICANT: Algate, Paul A.  
;; APPLICANT: Secrist, Heather  
;; APPLICANT: Indrias, Carol Yoseph  
;; APPLICANT: Benson, Darin R.  
;; APPLICANT: Elliott, Mark  
;; APPLICANT: Mannion, Jane  
;; APPLICANT: Kalos, Michael D.  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
;; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
;; FILE REFERENCE: 210121.475C9  
;; CURRENT APPLICATION NUMBER: US/09/738,973  
;; NUMBER OF SEQ ID NOS: 587  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 442  
;; LENGTH: 337  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-738-973-442

## Alignment Scores:

Pred. No.: 3.54e-94 Length: 337  
Score: 97.00 Matches: 97

Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0

US-09-854-133-586 (1-97) x US-09-738-973-442 (1-337)

QY 1 GluValGluValSerArgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThr 20  
Db 5 GAGTTGAAGTGAGCAGAGATCATGCCCTGGGTGACAGTGAAGACTCTGTCTCAACA 64  
QY 21 GluLeuArgLysLysGluArgLysLysArgGluArgLysPheGlnAlaAsnCysGly 40  
Db 65 GAATTAAGGAAAAAGAAAGAAAAAGAGAGAGAGAAATCCAGGCCAATTGTGGC 124  
QY 41 IleAspPheIleIlePheTrpIlePheTrpIleLeuLeuPheSerHisTrpIleGln 60  
Db 125 ATAGATTATCATATCTGTGATTTTGGATTCTTTGTTCTCATCACTGATTTCAG 184  
QY 61 GluSerLeuLeuCysProProSerProLysGluValThrCysArgGluMetLeuThrGly 80  
Db 185 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGAGTTACTGCAGGGAATGTAAACGGGA 244  
QY 81 GlyCysLeuProTrpAlaThrArgSerHisLeuGlyArgArgLysCysSer 97  
Db 245 GGCTGCCTTCCCTGGGCAACAAGAGCCACTGGGCGAGAGAAAGTGACGC 295

## RESULT 3

US-09-854-133-441

;; Sequence 441, Application US/09854133  
;; Publication No. US20020183499A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Lodes, Michael J.  
;; APPLICANT: Mohamath, Raodoh  
;; APPLICANT: Henderson, Robert A.  
;; APPLICANT: Benson, Darin R.  
;; APPLICANT: Secrist, Heather  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
;; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
;; FILE REFERENCE: 210121.475C10  
;; CURRENT APPLICATION NUMBER: US/09/854,133  
;; NUMBER OF SEQ ID NOS: 735  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 441  
;; LENGTH: 5981  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-854-133-441

## Alignment Scores:

Pred. No.: 5.6e-92 Length: 5981  
Score: 96.00 Matches: 96  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.97% Indels: 0  
DB: Gaps: 0

US-09-854-133-586 (1-97) x US-09-854-133-441 (1-5981)

QY 2 ValGluValSerArgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThr 21  
Db 3 GTTGAAGTGAGCAGAGATCATGCCCTGGGTGACAGTGAAGACTCTGTCTCAACA 62  
QY 22 LeuArgLysLysGluArgLysLysArgGluArgLysPheGlnAlaAsnCysGly 41  
Db 63 TTAAGGAAAAAGAAAGAAAAAGAGAGAGAGAAATTCAGGCCAATTGTGGCATA 122  
QY 42 AspPheIleIlePheTrpIlePheTrpIleLeuLeuPheSerHisTrpIleGln 61  
Db 123 GATTTATCATATCTGTGATTTTGGATTCTTTGTTCTCATCACTGATTTCAGGAA 182  
QY 62 SerLeuLeuCysProProSerProLysGluValThrCysArgGluMetLeuThrGly 81

Db 183 AGCCTGTGTGTCCACCATCTCCAAAGAGGTACCTGCAGGGAATGTTAACGGAGGC 242

QY 82 CysLeuProTPrAlaThrArgSerHisLeuGlyArgArgLysCysSer 97  
Db 243 TGCCTTCCCTGGGCAACAGAGCCACCTGGGAGAGAAAGTGCAGC 290

## RESULT 4

US-09-738-973-441  
Sequence 441, Application US/09738973  
Patent No. US20020110563A1

## GENERAL INFORMATION:

APPLICANT: Reed, Steven G.  
APPLICANT: Henderson, Robert A.  
APPLICANT: Lodes, Michael J.  
APPLICANT: Fling, Steven P.  
APPLICANT: Mohamath, Raodoh  
APPLICANT: Algate, Paul A.  
APPLICANT: Secrist, Heather  
APPLICANT: Indrias, Carol Yoseph  
APPLICANT: Benson, Darin R.  
APPLICANT: Elliot, Mark  
APPLICANT: Mannion, Jane  
APPLICANT: Kalos, Michael D.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.475C9  
CURRENT APPLICATION NUMBER: US/09/738, 973  
CURRENT FILING DATE: 2000-12-14  
NUMBER OF SEQ ID NOS: 587  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 441  
LENGTH: 5981  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-738-973-441

## Alignment Scores:

Pred. No.: 5.6e-92 Length: 5981  
Score: 96.00 Matches: 96  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.97% Indels: 0  
DB: 10 Gaps: 0

US-09-854-133-586 (1-97) x US-09-738-973-441 (1-5981)

QY 2 ValGluValSerArgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThrGlu 21

Db 3 GTTGAAGTGAAGAGAGATCATGCGACCTGGGTGACAGTGAAGTCTGTCTCAACAGAA 62

QY 22 LeuArgLysLysGluArgLysLysLysArgGluArgLysPheGlnAlaAsnCysGlyLe 41

Db 63 TTAAGGAAAAAAGAAAGAAAGAAAGAGAGAAATTCAGGCCCAATTGTGCATA 122

QY 42 AspPheIleIlePheTrpIlePheTrpIleLeuLeuPheSerHisHisTrpIleGlnGlu 61

Db 123 GATTATATCATATTTCTGATTTTGTGATTTCTTTGTTCTCAATGATTCAGGAA 182

QY 62 SerLeuLeuCysProProSerProLysGluValThrCysArgGluMetLeuThrGlyGly 81

Db 183 AGCCTGTGTGTCCACCATCTCCAAAGAGGTTACCTGCAGGGAATGTTAAAGGAGGC 242

QY 82 CysLeuProTPrAlaThrArgSerHisLeuGlyArgArgLysCysSer 97

Db 243 TGCCTTCCCTGGGCAACAGAGCCACCTGGGAGAGAAAGTGCAGC 290

## RESULT 5

US-09-854-133-440  
Sequence 440, Application US/09854133  
Publication No. US20020183499A1

GENERAL INFORMATION:  
APPLICANT: Lodes, Michael J.

APPLICANT: Mohamath, Raodoh  
APPLICANT: Henderson, Robert A.  
APPLICANT: Benson, Darin R.

APPLICANT: Secrist, Heather

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR

TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.475C10

CURRENT APPLICATION NUMBER: US/09/854, 133

CURRENT FILING DATE: 2001-05-11

NUMBER OF SEQ ID NOS: 735

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 440

LENGTH: 2239

TYPE: DNA

ORGANISM: Homo sapiens

US-09-854-133-440

## Alignment Scores:

Pred. No.: 3.45e-89 Length: 2239  
Score: 93.00 Matches: 93  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 95.88% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-09-854-133-440 (1-2239)

QY 1 GluValGluValSerArgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThr 20

Db 2 GAGGTGAAGTGAAGAGAGATCATGCGACCTGGGTGACAGTGAAGTCTGTCTCAACA 61

QY 21 GluLeuArgLysLysGluArgLysLysLysArgGluArgLysPheGlnAlaAsnCysGly 40

Db 62 GAATTAAGGAAAAAAGAAAGAAAGAAAGAGAGAGAGAAATTCAGGCCAATTGTGCG 121

QY 41 IleAspPheIleIlePheTrpIlePheTrpIleLeuLeuPheSerHisHisTrpIleGln 60

Db 122 ATGATTTTATCATATTTCTGATTTTGTGATTTCTTTCTCATCATGATTCAG 181

QY 61 GluSerLeuLeuCysProProSerProLysGluValThrCysArgGluMetLeuThrGly 80

Db 182 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGGTTACCTGCAGGGAATGTTAAAGGGA 241

QY 81 GlyCysLeuProTPrAlaThrArgSerHisLeuGlyArg 93

Db 242 GGCTGCTTCCCTGGGCAACAGAGCCACCTGGGAGG 280

## RESULT 6

US-09-738-973-440  
Sequence 440, Application US/09738973  
Patent No. US20020110563A1

## GENERAL INFORMATION:

APPLICANT: Reed, Steven G.  
APPLICANT: Henderson, Robert A.  
APPLICANT: Lodes, Michael J.

APPLICANT: Fling, Steven P.

APPLICANT: Mohamath, Raodoh

APPLICANT: Algate, Paul A.

APPLICANT: Secrist, Heather

APPLICANT: Indrias, Carol Yoseph

APPLICANT: Benson, Darin R.

APPLICANT: Elliot, Mark

APPLICANT: Mannion, Jane

APPLICANT: Kalos, Michael D.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR

TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.475C9

CURRENT APPLICATION NUMBER: US/09/738, 973

CURRENT FILING DATE: 2000-12-14

NUMBER OF SEQ ID NOS: 587

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 440

LENGTH: 2239

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-738-973-440

```

**Alignment Scores:**

File: NO.:	3.45e-89	Length:	2239
Score:	93.00	Matches:	93
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	95.88%	Indels:	0
DB:	10	Gaps:	0

US-09-854-133-586 (1-97) x US-09-738-973-440 (1-2239)

QY	1	GlValGluValSerArgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThr	20
Db	2	GAGTTGAAGTGAAGCAGACAGATCATGCGCCCTGGGTGACAGTGAAGACTCTGTCTCAACA	61
QY	21	GluLeuArgLysLysGluArgLysLysLysArgGluArgLysPheGlnAlaAsnCysGly	40
Db	62	GATTTAAGCAAAAAAGAAAGAAAAAGAGAGAGAGAAATTCAGGCCAATTGTGGC	121
QY	41	IleAspPheIleIlePheTrpIlePheTrpIleLeuLeuPheSerHisIleTrpIleGln	60
Db	122	ATAGATTTTATCATATTCTGGATTCTTTGGATTCTTTGTTTCTCATCTGACATTTCAG	181
QY	61	GluSerLeuLeuCysProProSerProLysGluValThrCysArgGluMetLeuThrGly	80
Db	182	GAAAGCCTGTTGTGTCCACCACCATCTCCAAAGAGAGTTACCTGCAGGGAATGTTTAACGGGA	241
QY	81	GlyCysLeuProTrpAlaThrArgSerHisLeuGlyArg	93
Db	242	GGCTGCTTCCCTGGGCAACAAGAGACCACTGGGCAGG	280

## RESULT 7

```

US-10-163-866-32
; Sequence 32, Application US/10163866
; Publication No. US20030027188A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-080C
; CURRENT APPLICATION NUMBER: US/10/163, 866
; CURRENT FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: US 60/296, 076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328, 605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/338, 733
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/357, 253
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: US 60/357, 600
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32
; LENGTH: 520
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-163-866-32

```

Alignment Scores:

Pred. No.:	1.98e-30	Length:	520
Score:	37.00	Matches:	37
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	38.14%	Indels:	0
DB:	9	Gaps:	0

US-09-854-133-586 (1-97) x US-10-163-866-32 (1-520)

QY 61 GluSerLeuLeuGlyProProSerProGlyGluValThrCysArgGluMetLeuThrGly 80

Db 8 GAAAGCCTGTTGTCACCATCTCCAAAGAGAGTTACTGACGGAATGTTAACGGA 67

QY 81 GlyCysLeuProT<sup>r</sup>Palat<sup>r</sup>ArgSerHisLeuGlyArgArgLysCysSer 97  
 Db 68 GCCTGCCTTCCCTGGGCAACAAGAGGCCACCTGGGCAGAGAGAAGTGCAC 118

## RESULT 8

```

US-10-046-935-1307
; Sequence 1307, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1307
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9, 19, 461, 497, 500, 502
; OTHER INFORMATION: n = A,T,C or G
US-10-046-935-1307

```

Alignment Scores:

```

pred. No.:      2,16e-30      length:      572
Score:          37.00        Matches:      37
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match:     38.14%       Indels:       0
DB:              9           Gaps:       0

```

US-09-854-133-586 (1-97) x US-10-046-935-1307 (1-572)

QY		61	G l u s e r l e u L e u C y s P r o P r o S e r P r o L y g l u V a I T h r C y s a r g G l u m e t l e u t h r g l y	80
D b		256	G A A G C C T G T T G T G T C A C C A C T C T C A A A G A G G T T A C C T G C A G G A A T G T T A A C G G A	315
QY		81	G l y c y s l e u P r o f i p a l a t h r a r s e r H i s l e u g l y a r g a r g l y c y s s e r	97
D b		316	G G C T G C C T T C C C T G G G C A C A A G A G A C C A C C T G G G C A G A G A A A G T C A G C	366

## RESULT 9

```

US-09-878178-1307
; Sequence 1307, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527
; CURRENT APPLICATION NUMBER: US/09/878,178
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 2237
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1307
;
; LENGTH: 572
;
; TYPE: DNA
;
; ORGANISM: Homo sapien
;
; FEATURE:
;
; NAME/KEY: misc_feature

```



LOCATION: (1)...(572)  
OTHER INFORMATION: n = A,T,C or G  
US-09-878-178-1307

## Alignment Scores:

Pred. No.:	2.16e-30	Length:	572
Score:	37.00	Matches:	37
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	38.14%	Indels:	0
DB:	9	Gaps:	0

US-09-854-133-586 (1-97) x US-09-878-178-1307 (1-572)

QY 61 GluSerLeuLeuCyProProSerProLysGluValThrCysArgGluMetLeuThrGly 80  
|||||  
DB 256 GAAAGCCTGTGTGTCACCATCTCCAAAGAGGTTACCTGCAGGGAATGTTAAGGGA 315  
QY 81 GlyCysLeuProTTPalaThrArgSerHisLeuGlyArgArgLysCysSer 97  
|||||  
DB 316 GGCTGCCTTCCCTGGGCAACAAGAGCCACCTGGGCGAGAGAAAGTGACAGC 366

## RESULT 10

US-10-146-502-1307  
Sequence 1307, Application US/10146502  
Publication No. US20030069180A1  
GENERAL INFORMATION:  
APPLICANT: Jiang, Yugu  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Secrist, Heather  
APPLICANT: Wang, Aijun  
APPLICANT: Stolk, John A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER  
FILE REFERENCE: 210121.527C2  
CURRENT APPLICATION NUMBER: US/10/146,502  
CURRENT FILING DATE: 2002-05-14  
NUMBER OF SEQ ID NOS: 2241  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1307  
LENGTH: 572  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 9, 19, 461, 497, 500, 502  
OTHER INFORMATION: n = A,T,C or G  
US-10-146-502-1307

## Alignment Scores:

Pred. No.:	2.16e-30	Length:	572
Score:	37.00	Matches:	37
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	38.14%	Indels:	0
DB:	9	Gaps:	0

US-09-854-133-586 (1-97) x US-10-146-502-1307 (1-572)

QY 61 GluSerLeuLeuCyProProSerProLysGluValThrCysArgGluMetLeuThrGly 80  
|||||  
DB 256 GAAAGCCTGTGTGTCACCATCTCCAAAGAGGTTACCTGCAGGGAATGTTAAGGGA 315  
QY 81 GlyCysLeuProTTPalaThrArgSerHisLeuGlyArgArgLysCysSer 97  
|||||  
DB 316 GGCTGCCTTCCCTGGGCAACAAGAGCCACCTGGGCGAGAGAAAGTGACAGC 366

## RESULT 11

US-10-163-866-53  
Sequence 53, Application US/10163866  
Publication No. US20030027188A1  
GENERAL INFORMATION:  
APPLICANT: EXELIXIS, INC.

TITLE OF INVENTION: SICTS AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE  
FILE REFERENCE: EX02-080C  
CURRENT APPLICATION NUMBER: US/10/163,866  
CURRENT FILING DATE: 2002-06-05  
PRIOR APPLICATION NUMBER: US 60/296,076  
PRIOR FILING DATE: 2001-06-05  
PRIOR APPLICATION NUMBER: US 60/328,605  
PRIOR FILING DATE: 2001-10-10  
PRIOR APPLICATION NUMBER: US 60/338,733  
PRIOR FILING DATE: 2001-10-22  
PRIOR APPLICATION NUMBER: US 60/357,253  
PRIOR FILING DATE: 2002-02-15  
PRIOR APPLICATION NUMBER: US 60/357,600  
PRIOR FILING DATE: 2002-02-15  
NUMBER OF SEQ ID NOS: 54  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 53  
LENGTH: 1268  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-163-866-53

## Alignment Scores:

Pred. No.:	4.46e-30	Length:	1268
Score:	37.00	Matches:	37
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	38.14%	Indels:	0
DB:	9	Gaps:	0

US-09-854-133-586 (1-97) x US-10-163-866-53 (1-1268)

QY 61 GluSerLeuLeuCyProProSerProLysGluValThrCysArgGluMetLeuThrGly 80  
|||||  
DB 30 GAAAGCCTGTGTGTCACCATCTCCAAAGAGGTTACCTGCAGGGAATGTTAAGGGA 89  
QY 81 GlyCysLeuProTTPalaThrArgSerHisLeuGlyArgArgLysCysSer 97  
|||||  
DB 90 GGCTGCCTTCCCTGGGCAACAAGAGCCACCTGGGCGAGAGAAAGTGACAGC 140

## RESULT 12

US-10-163-866-52  
Sequence 52, Application US/10163866  
Publication No. US20030027188A1  
GENERAL INFORMATION:  
APPLICANT: EXELIXIS, INC.  
TITLE OF INVENTION: SICTS AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE  
FILE REFERENCE: EX02-080C  
CURRENT APPLICATION NUMBER: US/10/163,866  
CURRENT FILING DATE: 2002-06-05  
PRIOR APPLICATION NUMBER: US 60/296,076  
PRIOR FILING DATE: 2001-06-05  
PRIOR APPLICATION NUMBER: US 60/328,605  
PRIOR FILING DATE: 2001-10-10  
PRIOR APPLICATION NUMBER: US 60/338,733  
PRIOR FILING DATE: 2001-10-22  
PRIOR APPLICATION NUMBER: US 60/357,253  
PRIOR FILING DATE: 2002-02-15  
PRIOR APPLICATION NUMBER: US 60/357,600  
PRIOR FILING DATE: 2002-02-15  
NUMBER OF SEQ ID NOS: 54  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 52  
LENGTH: 1528  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-163-866-52

## Alignment Scores:

Pred. No.:	5.28e-30	Length:	1528
Score:	37.00	Matches:	37
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0

Query Match: 38.14% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-163-866-52 (1-1528)

QY 61 GluSerLeuLeuCySProProSerProLysGluValThrCysArgGluMetLeuThrgly 80  
DB 30 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGGTTACCTGCAGGGAATGTTAACGGGA 89

QY 81 GlyCysLeuProTrrPalatThrArgSerHisLeuGlyArgArgLysCysSer 97  
DB 90 GGCTGCCTTCCCTGGGCAACAAGAGCCACCTGGGAGAGAAAGTGCAGC 140

RESULT 13  
US-10-163-866-33

; Sequence 33, Application US/10163866  
; Publication No. US20030027188A1  
; GENERAL INFORMATION:

; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX02-080C  
; CURRENT APPLICATION NUMBER: US/10/163,866  
; PRIOR APPLICATION NUMBER: US 60/296,076  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/328,605  
; PRIOR FILING DATE: 2001-10-10  
; PRIOR APPLICATION NUMBER: US 60/338,733  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/357,253  
; PRIOR FILING DATE: 2002-02-15  
; PRIOR APPLICATION NUMBER: US 60/357,600  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 33  
; LENGTH: 1542  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-163-866-33

Alignment Scores:  
Pred. No.: 5.33e-30 Length: 1542  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.14% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-163-866-33 (1-1542)

QY 61 GluSerLeuLeuCySProProSerProLysGluValThrCysArgGluMetLeuThrgly 80  
DB 16 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGGTTACCTGCAGGGAATGTTAACGGGA 75

QY 81 GlyCysLeuProTrrPalatThrArgSerHisLeuGlyArgArgLysCysSer 97  
DB 76 GGCTGCCTTCCCTGGGCAACAAGAGCCACCTGGGAGAGAAAGTGCAGC 126

RESULT 14  
US-10-163-866-30

; Sequence 30, Application US/10163866  
; Publication No. US20030027188A1  
; GENERAL INFORMATION:

; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX02-080C  
; CURRENT APPLICATION NUMBER: US/10/163,866  
; PRIOR APPLICATION NUMBER: US 60/296,076  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/328,605  
; PRIOR FILING DATE: 2001-10-10

; SEQ ID NO 30  
; LENGTH: 1542  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-163-866-30

Alignment Scores:  
Pred. No.: 6.76e-30 Length: 2000  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.14% Indels: 0  
DB: 9 Gaps: 0

; PRIOR APPLICATION NUMBER: US 60/338,733  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/357,253  
; PRIOR FILING DATE: 2002-02-15  
; PRIOR APPLICATION NUMBER: US 60/357,600  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 30  
; LENGTH: 1861  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-163-866-30

Alignment Scores:

Pred. No.: 6.33e-30 Length: 1861  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.14% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-163-866-30 (1-1861)

QY 61 GluSerLeuLeuCySProProSerProLysGluValThrCysArgGluMetLeuThrgly 80  
DB 243 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGGTTACCTGCAGGGAATGTTAACGGGA 302

QY 81 GlyCysLeuProTrrPalatThrArgSerHisLeuGlyArgArgLysCysSer 97  
DB 303 GGCTGCCTTCCCTGGGCAACAAGAGCCACCTGGGAGAGAAAGTGCAGC 353

RESULT 15  
US-10-163-866-34

; Sequence 34, Application US/10163866  
; Publication No. US20030027188A1  
; GENERAL INFORMATION:

; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX02-080C  
; CURRENT APPLICATION NUMBER: US/10/163,866  
; PRIOR APPLICATION NUMBER: US 60/296,076  
; PRIOR FILING DATE: 2002-06-05  
; PRIOR APPLICATION NUMBER: US 60/328,605  
; PRIOR FILING DATE: 2001-10-10  
; PRIOR APPLICATION NUMBER: US 60/338,733  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/357,253  
; PRIOR FILING DATE: 2002-02-15  
; PRIOR APPLICATION NUMBER: US 60/357,600  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 34  
; LENGTH: 2000  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-163-866-34

Alignment Scores:

Pred. No.: 6.76e-30 Length: 2000  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.14% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-163-866-34 (1-2000)

QY 61 GluSerLeuLeuCySProProSerProLysGluValThrCysArgGluMetLeuThrgly 80  
DB 143 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGGTTACCTGCAGGGAATGTTAACGGGA 202

OY 81 GLYCISLEUPROTRPALATHRARGSERHISLEUGLYARGGLYSCYSSER 97  
DB 203 GGCTGCCCTTCCCTGGCAACAAGAGCACCCTGGCAGAGAAAGTGACGC 253

RESULT 16  
US-10-163-866-29

; Sequence 29, Application US/10163866  
; Publication No. US20030027188A1  
; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: SLCS AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX02-080C  
; CURRENT APPLICATION NUMBER: US/10/163,866  
; PRIOR FILING DATE: 2002-06-05  
; PRIOR APPLICATION NUMBER: US 60/296,076  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/328,605  
; PRIOR FILING DATE: 2001-10-10  
; PRIOR APPLICATION NUMBER: US 60/338,733  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/357,253  
; PRIOR FILING DATE: 2002-02-15  
; PRIOR APPLICATION NUMBER: US 60/357,600  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 29  
; LENGTH: 2482  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-163-866-29

Alignment Scores:  
Pred. No.: 8.23e-30 Length: 2482  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.14% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-163-866-29 (1-2482)

OY 61 GLUSERLEUCYSPROPROSERPROLYSGLUVALTHRCYSARGGLUMETLEUTHRGly 80  
DB 239 GAAAGCCTGTGTGTCACCATCTCCAAAGAGGTTTACCTGCAGGGAATGTTAACGGGA 298

OY 81 GLYCISLEUPROTRPALATHRARGSERHISLEUGLYARGGLYSCYSSER 97  
DB 299 GGCTGCCCTTCCCTGGCAACAAGAGCACCCTGGCAGAGAAAGTGACGC 349

## RESULT 17

US-10-163-866-31  
; Sequence 31, Application US/10163866  
; Publication No. US20030027188A1  
; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: SLCS AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX02-080C  
; CURRENT APPLICATION NUMBER: US/10/163,866  
; PRIOR FILING DATE: 2002-06-05  
; PRIOR APPLICATION NUMBER: US 60/296,076  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/328,605  
; PRIOR FILING DATE: 2001-10-10  
; PRIOR APPLICATION NUMBER: US 60/338,733  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/357,253  
; PRIOR FILING DATE: 2002-02-15  
; PRIOR APPLICATION NUMBER: US 60/357,600  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 31  
; LENGTH: 3144  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-163-866-31

Alignment Scores:  
Pred. No.: 1.02e-29 Length: 3144  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.14% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-163-866-31 (1-3144)

OY 61 GLUSERLEUCYSPROPROSERPROLYSGLUVALTHRCYSARGGLUMETLEUTHRGly 80  
DB 8 GAAAGCCTGTGTGTCACCATCTCCAAAGAGGTTTACCTGCAGGGAATGTTAACGGGA 67

OY 81 GLYCISLEUPROTRPALATHRARGSERHISLEUGLYARGGLYSCYSSER 97  
DB 68 GGCTGCCCTTCCCTGGCAACAAGAGCACCCTGGCAGAGAAAGTGACGC 118

## RESULT 18

US-09-764-877-3626  
; Sequence 3626, Application US/09764877  
; Patent No. US20020147140A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005  
; CURRENT APPLICATION NUMBER: US/09/764,877  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: US 60/357,600  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3626  
; LENGTH: 29607  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-877-3626

Alignment Scores:  
Pred. No.: 0.0244 Length: 29607  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
DB: 10 Gaps: 0

US-09-854-133-586 (1-97) x US-09-764-877-3626 (1-29607)

OY 8 HISALASERLEUGLYASPSERGLUTHRLEUSERGLN 19  
DB 24558 CACGCTAGCCTGGGTGACAGTACAGCCCTGTCTCAA 24593

## RESULT 19

US-09-933-797-216  
; Sequence 216, Application US/09933797  
; Patent No. US20020155119A1  
; GENERAL INFORMATION:  
; APPLICANT: Robert A. Sikes et al.  
; TITLE OF INVENTION: Isolation and Use of Fetal Drogenital  
; FILE REFERENCE: 9901-007-999  
; CURRENT APPLICATION NUMBER: US/09/933,797  
; PRIOR FILING DATE: 2001-08-22  
; PRIOR APPLICATION NUMBER: US/09/482,933  
; PRIOR FILING DATE: 2000-01-14  
; PRIOR APPLICATION NUMBER: PCT/US99/10746  
; PRIOR FILING DATE: 1999-05/14  
; PRIOR APPLICATION NUMBER: 60/085,383

; PRIOR FILING DATE: 1998-05-14  
; NUMBER OF SEQ ID NOS: 811  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 216  
; LENGTH: 294  
; TYPE: DNA  
; ORGANISM: Murine  
; FEATURE:  
; NAME/KEY: misc-feature  
; LOCATION: (1)...(294)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-933-797-216

## Alignment Scores:

Pred. No.:	0.00416	Length:	294
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	11.34%	Indels:	0
DB:	9	Gaps:	0

US-09-854-133-586 (1-97) x US-09-933-797-216 (1-294)

QY 23 ArglyslYsgluArglyslYsgluArg 33  
Db 61 AGAAAGAAAGAAAGAAAGAAAGAGAGAGA 93

## RESULT 20

US-10-091-504-2209  
; Sequence 2209, Application US/10091504  
; Publication No. US20030059908A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC007C1  
; CURRENT APPLICATION NUMBER: US/10/091,504  
; CURRENT FILING DATE: 2002-03-07  
; NUMBER OF SEQ ID NOS: 2442  
; Prior Application removed - See File Wrapper or Palm  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2209  
; LENGTH: 32190  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-091-504-2209

## Alignment Scores:

Pred. No.:	0.302	Length:	32190
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	11.34%	Indels:	0
DB:	9	Gaps:	0

US-09-854-133-586 (1-97) x US-10-091-504-2209 (1-32190)

QY 9 AlaserleuglyaspserglutThrleuSergln 19  
Db 30055 GCCAGCCTGGGGCAGACGAGACTCTGTCTCAA 30087

## RESULT 21

US-09-764-869-2209  
; Sequence 2209, Application US/09764869  
; Patent No. US20020061521A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC007  
; CURRENT APPLICATION NUMBER: US/09/764,869  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 2442  
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2209  
; LENGTH: 32190  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-764-869-2209

## Alignment Scores:

Pred. No.:	0.302	Length:	32190
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	11.34%	Indels:	0
DB:	10	Gaps:	0

US-09-854-133-586 (1-97) x US-09-764-869-2209 (1-32190)

QY 9 AlaserleuglyaspserglutThrleuSergln 19  
Db 30055 GCCAGCCTGGGGCAGACGAGACTCTGTCTCAA 30087

## RESULT 22

US-09-998-598-2201/c  
; Sequence 2201, Application US/09998598  
; Patent No. US20020150922A1  
; GENERAL INFORMATION:  
; APPLICANT: Stolk, John A.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Chenault, Ruth A.  
; APPLICANT: Meagher, Madelein Joy  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.561  
; CURRENT APPLICATION NUMBER: US/09/998,598  
; CURRENT FILING DATE: 2001-11-16  
; NUMBER OF SEQ ID NOS: 2606  
; SOFTWARE: Corixa Invention Disclosure Database  
; SEQ ID NO 2201  
; LENGTH: 261  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-998-598-2201

## Alignment Scores:

Pred. No.:	0.0428	Length:	261
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	10	Gaps:	0

US-09-854-133-586 (1-97) x US-09-998-598-2201 (1-261)

QY 10 SerleuglyaspserglutThrleuSergln 19  
Db 104 AGCCTGGGTGACAGTGAAGACTCTGTCTCAA 75

## RESULT 23

US-09-803-719-71  
; Sequence 71, Application US/09803719  
; Publication No. US20030044783A1  
; GENERAL INFORMATION:  
; APPLICANT: Williams, Lewis T.  
; APPLICANT: Escobedo, Jaime  
; APPLICANT: Innis, Michael A.  
; APPLICANT: Garcia, Pablo Dominguez  
; APPLICANT: Sudduth-Klinger, Julie  
; APPLICANT: Reinhard, Christoph  
; APPLICANT: Glese, Klaus  
; APPLICANT: Randazzo, Filippo  
; APPLICANT: Kennedy, Giulia C.  
; APPLICANT: Pot, David  
; APPLICANT: Kassam, Altaf  
; APPLICANT: Lamson, George



```
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, Lee William
; APPLICANT: Stache-Crain, Birgit
; TITLE OF INVENTION: Human Genes and Gene Products
; FILE REFERENCE: 1624.002
; CURRENT APPLICATION NUMBER: US/09/803,719
; CURRENT FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/188,609
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 2396
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-803-719-71
```

```
Alignment Scores:
Pred. No.: 0.0473 Length: 291
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.31% Indels: 0
DB: 9 Gaps: 0
```

US-09-854-133-586 (1-97) x US-09-803-719-71 (1-291)

QY 10 SerleuGlyAspSerGluThrLeuSergln 19

DB 97 AGCCTGGGTGACAGCGAGACTCTATCTCAA 126

#### RESULT 24

US-09-803-719-234

```
; Sequence 234, Application US/09803719
; Publication No. US20030044783A1
```

#### GENERAL INFORMATION:

```
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominiguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pot, David
; APPLICANT: Kassam, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, Lee William
; APPLICANT: Stache-Crain, Birgit
; TITLE OF INVENTION: Human Genes and Gene Products
; FILE REFERENCE: 1624.002
; CURRENT APPLICATION NUMBER: US/09/803,719
; CURRENT FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/188,609
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 2396
; SOFTWARE: FastSeq for Windows Version 4.0
```

SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 234
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-803-719-234
```

```
Alignment Scores:
Pred. No.: 0.0473 Length: 291
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.31% Indels: 0
DB: 9 Gaps: 0
```

US-09-854-133-586 (1-97) x US-09-803-719-234 (1-291)

QY 10 SerleuGlyAspSerGluThrLeuSergln 19

DB 28 AGCCTGGGTGACAGCGAGACTCTATCTCAA 57

#### RESULT 25

US-09-803-719-83

```
; Sequence 83, Application US/09803719
; Publication No. US20030044783A1
```

#### GENERAL INFORMATION:

```
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominiguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulia C.
; APPLICANT: Pot, David
; APPLICANT: Kassam, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, Lee William
; APPLICANT: Stache-Crain, Birgit
; TITLE OF INVENTION: Human Genes and Gene Products
; FILE REFERENCE: 1624.002
; CURRENT APPLICATION NUMBER: US/09/803,719
; CURRENT FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/188,609
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 2396
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 83
; LENGTH: 303
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-803-719-83
```

```
Alignment Scores:
Pred. No.: 0.049 Length: 303
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.31% Indels: 0
DB: 9 Gaps: 0
```

US-09-854-133-586 (1-97) x US-09-803-719-83 (1-303)

QY 10 SerleuGlyAspSerGluThrLeuSergln 19

DB 28 AGCCTGGGTGACAGCGAGACTCTATCTCAA 57

Db 97 AGCCTGGGTGACAGCAGACTCTATCTCAA 126

## RESULT 26

US-09-954-531-403/c  
; Sequence 403, Application US/09954531  
; Patent No. US20020165180A1  
; GENERAL INFORMATION:

; APPLICANT: Weaver, Zoe

; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc

; FILE REFERENCE: Gene Sets

; CURRENT APPLICATION NUMBER: US/09/954,531

; CURRENT FILING DATE: 2002-05-02

; PRIOR APPLICATION NUMBER: US/60/233,133

; PRIOR FILING DATE: 2000-09-18

; PRIOR APPLICATION NUMBER: US/60/234,009

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: US/60/234,034

; PRIOR FILING DATE: 2000-09-20

; PRIOR APPLICATION NUMBER: US/60/234,509

; PRIOR FILING DATE: 2000-09-22

; PRIOR APPLICATION NUMBER: US/60/234,567

; NUMBER OF SEQ ID NOS: 1392

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 403

; LENGTH: 351

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(351)

; OTHER INFORMATION: n=a,t,g or c

US-09-954-531-403

## Alignment Scores:

Pred. No.:	0.0561	Length:	351
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	9	Gaps:	0

US-09-854-133-586 (1-97) x US-09-954-531-403 (1-351)

QY 10 SerLeuGlyAspSergIuThrLeuSergIn 19

Db 153 AGCCTGGGTGACAGCAGACCTGTCTCAA 124

## RESULT 27

US-09-867-701-7120/c

; Sequence 7120, Application US/09867701

; Patent No. US20020132237A1

; GENERAL INFORMATION:

; APPLICANT: Aglate, Paul A.

; APPLICANT: Jones, Robert

; APPLICANT: Harlocker, Susan L.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.497

; CURRENT APPLICATION NUMBER: US/09/867,701

; CURRENT FILING DATE: 2001-05-29

; NUMBER OF SEQ ID NOS: 10912

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7120

; LENGTH: 418

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-867-701-7120

## Alignment Scores:

Pred. No.:	0.0658	Length:	418
Score:	10.00	Matches:	10

Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	10	Gaps:	0

US-09-854-133-586 (1-97) x US-09-867-701-7120 (1-418)

QY 10 SerLeuGlyAspSergIuThrLeuSergIn 19

Db 336 AGCCTGGGACAGCAGACTCTGTCTCAA 307

## RESULT 28

US-09-918-995-12641

; Sequence 12641, Application US/09918995

; Publication No. US20030073623A1

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

; FILE REFERENCE: 20411-756

; CURRENT APPLICATION NUMBER: US/09/918,995

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: US/09/235,076

; PRIOR FILING DATE: 1999-01-20

; NUMBER OF SEQ ID NOS: 38054

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 12641

; LENGTH: 483

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(483)

; OTHER INFORMATION: n = A,T,C or G

US-09-918-995-12641

## Alignment Scores:

Pred. No.:	0.0751	Length:	483
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	9	Gaps:	0

US-09-854-133-586 (1-97) x US-09-918-995-12641 (1-483)

QY 10 SerLeuGlyAspSergIuThrLeuSergIn 19

Db 423 AGCCTGGGCGACAGTGTGACTCTATCTCAA 452

## RESULT 29

US-09-764-891-1995/c

; Sequence 1995, Application US/09764891

; Publication No. US20030077808A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PC006

; CURRENT APPLICATION NUMBER: US/09/764,891

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 10231

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1995

; LENGTH: 497

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (109)

; OTHER INFORMATION: n equals a,t,g, or c

; NAME/KEY: SITE

; LOCATION: (489)

OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-891-1995

Alignment Scores:

Pred. No.:	0.077	Length:	497
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	9	Gaps:	0

US-09-854-133-586 (1-97) x US-09-764-891-1995 (1-497)

QY 10 Serleuglyaspserglutthreusergln 19  
Db 184 AGCCTGGGTGACAGTGAACCTGTCTCAA 155

RESULT 30  
US-09-864-761-12019/c  
; Sequence 12019, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 12019  
; LENGTH: 520  
; TYPE: DNA  
; ORGANISM: Homo sapiens

FEATURE:  
; OTHER INFORMATION: MAP TO AL136962.6  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.51  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.47  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.56  
US-09-864-761-12019

Alignment Scores:

Pred. No.:	0.0803	Length:	520
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	10	Gaps:	0

US-09-854-133-586 (1-97) x US-09-864-761-12019 (1-520)

QY 26 Gluarglysllysarggluarglysphe 35  
Db 262 GAAAGGAAAAAAGAGAGAGAGAAATTT 233

RESULT 31  
US-09-918-995-13252  
; Sequence 13252, Application US/09918995  
; Publication No. US20030073623A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCE: 20411-756  
; CURRENT APPLICATION NUMBER: US/09/918,995  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: US/09/235,076  
; PRIOR FILING DATE: 1999-01-20  
; NUMBER OF SEQ ID NOS: 38054  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 13252  
; LENGTH: 554  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(554)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-918-995-13252

Alignment Scores:

Pred. No.:	0.0851	Length:	554
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	9	Gaps:	0

US-09-854-133-586 (1-97) x US-09-918-995-13252 (1-554)

QY 10 Serleuglyaspserglutthreusergln 19  
Db 415 AGCCTGGGTGACAGCGAAGAACTGTCTCAA 444

RESULT 32  
US-10-028-072-255  
; Sequence 255, Application US/10028072  
; Publication No. US20030004311A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang

## FILE REFERENCE:

CURRENT APPLICATION NUMBER: US/10/028,072  
CURRENT FILING DATE: 2001-12-19  
PRIOR APPLICATION NUMBER: 60/049911  
PRIOR FILING DATE: 1997-06-18  
PRIOR APPLICATION NUMBER: 60/056974  
PRIOR FILING DATE: 1997-08-26  
PRIOR APPLICATION NUMBER: 60/059113  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059115  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059117  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059122  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059184  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059352  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059588  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059836  
PRIOR FILING DATE: 1997-09-24  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062285  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062287  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062814  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/062816  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063045  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063082  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/063127  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063327  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063329  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063550  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063561  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063704  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063733  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063735  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063738  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063755  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064248  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/064809  
PRIOR FILING DATE: 1997-11-07

PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065846  
PRIOR FILING DATE: 1997-11-17  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/066453  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066511  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/069212  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069278  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069334  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069694  
PRIOR FILING DATE: 1997-12-16  
PRIOR APPLICATION NUMBER: 60/072320  
PRIOR FILING DATE: 1998-01-23  
PRIOR APPLICATION NUMBER: 60/073612  
PRIOR FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: 60/074086  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/074092  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-02-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081695  
PRIOR FILING DATE: 1998-04-14  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081818  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082999  
PRIOR FILING DATE: 1998-04-24  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085149  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697



; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085704  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/086414  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/086430  
; PRIOR FILING DATE: 1998-05-22  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088730  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088741  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089947  
; PRIOR FILING DATE: 1998-06-19  
; PRIOR APPLICATION NUMBER: 60/090349  
; PRIOR FILING DATE: 1998-06-23  
; PRIOR APPLICATION NUMBER: 60/090429  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090445  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090538  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090863  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/091360  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091519  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091982  
; PRIOR FILING DATE: 1998-07-07

## Alignment Scores:

Pred. No.:	0.202	Length:	1432
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	9	Gaps:	0

US-09-854-133-586 (1-97) x US-10-028-072-255 (1-1432)

OY 10 Serleuglyaspsergluthrleusergl 19

DB 1261 AGCCTGGGTGACAGCGAGACCCTGTCTCAA 1290

## RESULT 33

US-10-121-049-255  
; Sequence 255, Application US/10121049  
; Publication No. US20030022239A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C17  
; CURRENT APPLICATION NUMBER: US/10/121,049  
; CURRENT FILING DATE: 2002-04-12  
; Prior Application removed - See file wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 255  
; LENGTH: 1432  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-121-049-255

## Alignment Scores:

Pred. No.:	0.202	Length:	1432
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	9	Gaps:	0

US-09-854-133-586 (1-97) x US-10-121-049-255 (1-1432)

OY 10 Serleuglyaspsergluthrleusergl 19

DB 1261 AGCCTGGGTGACAGCGAGACCCTGTCTCAA 1290

## RESULT 34

US-10-123-904-255  
; Sequence 255, Application US/10123904  
; Publication No. US20030022328A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C54  
; CURRENT APPLICATION NUMBER: US/10/123,904  
; CURRENT FILING DATE: 2002-04-16  
; Prior Application removed - See file wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 255  
; LENGTH: 1432  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-123-904-255

## Alignment Scores:

Pred. No.:	0.202	Length:	1432
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-123-904-255 (1-1432)

QY 10 SerleuglyAspSergluThrLeuSergln 19  
Db 1261 AGCCTGGGTGACAGCAGACCCCTGTCTCAA 1290

RESULT 35  
US-10-140-470-255

; Sequence 255, Application US/10140470  
; Publication No. US20030022331A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C160  
; CURRENT APPLICATION NUMBER: US/10/140,470  
; CURRENT FILING DATE: 2002-05-06  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 255  
; LENGTH: 1432  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-140-470-255

Alignment Scores:

Pred. No.: 0.202 Length: 1432  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-140-470-255 (1-1432)

QY 10 SerleuglyAspSergluThrLeuSergln 19

Db 1261 AGCCTGGGTGACAGCAGACCCCTGTCTCAA 1290

RESULT 36

US-10-175-746-255

; Sequence 255, Application US/10175746

; Publication No. US20030027270A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C353  
; CURRENT APPLICATION NUMBER: US/10/175,746  
; CURRENT FILING DATE: 2002-06-19  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 255  
; LENGTH: 1432  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-175-746-255

Alignment Scores:

Pred. No.: 0.202 Length: 1432  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-175-746-255 (1-1432)

QY 10 SerleuglyAspSergluThrLeuSergln 19

Db 1261 AGCCTGGGTGACAGCAGACCCCTGTCTCAA 1290

RESULT 37

US-10-176-918-255

; Sequence 255, Application US/10176918

; Publication No. US20030027275A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P3330R1C382  
; CURRENT APPLICATION NUMBER: US/10/176,918  
; CURRENT FILING DATE: 2002-06-20  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 255  
; LENGTH: 1432  
; TYPE: DNA  
; ORGANISM: Homo Sapien  
US-10-176-918-255

Alignment Scores:

Pred. No.: 0.202 Length: 1432  
Score: 10.00 Matches: 10

Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-176-921-255 (1-1432)

OY 10 Serleuglyaspsergluthrleusergln 19

DB 1261 AGCCTGGGTGACAGCAGACCCCTGTCTCAA 1290

## RESULT 38

US-10-176-921-255

; Sequence 255, Application US/10176921

; Publication No. US20030027276A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; TITLE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3330R1C288

; CURRENT APPLICATION NUMBER: US/10/176,921

; CURRENT FILING DATE: 2002-06-20

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 255

; LENGTH: 1432

; TYPE: DNA

; ORGANISM: Homo Sapien

; US-10-176-921-255

## Alignment Scores:

Pred. No.: 0.202

Score: 10.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 10.31%

DB: 9

US-09-854-133-586 (1-97) x US-10-176-921-255 (1-1432)

OY 10 Serleuglyaspsergluthrleusergln 19

DB 1261 AGCCTGGGTGACAGCAGACCCCTGTCTCAA 1290

## RESULT 39

US-10-137-865-255

; Sequence 255, Application US/10137865

; Publication No. US20030032155A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; TITLE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3330R1C154

; CURRENT APPLICATION NUMBER: US/10/137,865

; CURRENT FILING DATE: 2002-05-03

; Prior Application removed - See Palm or File Wrapper

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 255

; LENGTH: 1432

; TYPE: DNA

; ORGANISM: Homo Sapien

; US-10-137-865-255

## Alignment Scores:

Pred. No.: 0.202

Score: 10.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 10.31%

DB: 9

US-09-854-133-586 (1-97) x US-10-137-865-255 (1-1432)

OY 10 Serleuglyaspsergluthrleusergln 19

DB 1261 AGCCTGGGTGACAGCAGACCCCTGTCTCAA 1290

## RESULT 40

US-10-140-474-255

; Sequence 255, Application US/10140474

; Publication No. US20030032156A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; TITLE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3330R1C162

; CURRENT APPLICATION NUMBER: US/10/140,474

; CURRENT FILING DATE: 2002-05-06

; Prior Application removed - See Palm or File Wrapper

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 255

; LENGTH: 1432

; TYPE: DNA

; ORGANISM: Homo Sapien

; US-10-140-474-255

## Alignment Scores:

Pred. No.: 0.202

Score: 10.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 10.31%

DB: 9

US-09-854-133-586 (1-97) x US-10-137-865-255 (1-1432)

OY 10 Serleuglyaspsergluthrleusergln 19

DB 1261 AGCCTGGGTGACAGCAGACCCCTGTCTCAA 1290

Score: 10.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 10.31%  
DB: 9  
Matches: 10  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-10-140-474-255 (1-1432)

QY 10 SerleuglyAspsergluThrleusergln 19

Db 1261 AGCCTGGTGACAGCAGACCCCTGTCTCAA 1290

## RESULT 41

US-10-142-431-255

; Sequence 255, Application US/10142431

; Publication No. US20030036179A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: Deforge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3330R1C251

; CURRENT APPLICATION NUMBER: US/10/142,431

; CURRENT FILING DATE: 2002-05-10

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 255

; LENGTH: 1432

; TYPE: DNA

; ORGANISM: Homo Sapien

; US-10-142-431-255

## Alignment Scores:

Pred. No.: 0.202

Score: 10.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 10.31%

DB: 9

Length: 1432

Matches: 10

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-854-133-586 (1-97) x US-10-142-431-255 (1-1432)

QY 10 SerleuglyAspsergluThrleusergln 19

Db 1261 AGCCTGGTGACAGCAGACCCCTGTCTCAA 1290

## RESULT 42

US-10-143-114-255

; Sequence 255, Application US/10143114

; Publication No. US20030036180A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: Deforge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3330R1C211

; CURRENT APPLICATION NUMBER: US/10/143,114

; CURRENT FILING DATE: 2002-05-09

; Prior Application removed - See Palm or File Wrapper

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 255

; LENGTH: 1432

; TYPE: DNA

; ORGANISM: Homo Sapien

; US-10-143-114-255

## Alignment Scores:

Pred. No.: 0.202

Score: 10.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 10.31%

DB: 9

Length: 1432

Matches: 10

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-854-133-586 (1-97) x US-10-143-114-255 (1-1432)

QY 10 SerleuglyAspsergluThrleusergln 19

Db 1261 AGCCTGGTGACAGCAGACCCCTGTCTCAA 1290

## RESULT 43

US-10-140-002-255

; Sequence 255, Application US/10140002

; Publication No. US20030037623A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: Deforge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3330R1C59

; CURRENT APPLICATION NUMBER: US/10/140,002

; CURRENT FILING DATE: 2002-05-06

; Prior Application removed - See Palm or File Wrapper

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 255

; LENGTH: 1432

; TYPE: DNA

; ORGANISM: Homo Sapien

; US-10-140-002-255

## Alignment Scores:



Pred. No.: 0.202 Length: 1432  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-140-002-255 (1-1432)

QY 10 SerleuglyaspSergluThrluSergln 19

Db 1261 AGCCTGGGTGACAGCAGACCCCTGTCTCAA 1290

## RESULT 44

US-10-142-419-255

; Sequence 255, Application US/10142419

; Publication No. US20030044945A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; TITLE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3330R1C244

; CURRENT APPLICATION NUMBER: US/10/142,419

; CURRENT FILING DATE: 2002-05-10

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 255

; LENGTH: 1432

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-142-419-255

## Alignment Scores:

Pred. No.: 0.202 Length: 1432

Score: 10.00 Matches: 10

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 10.31% Indels: 0

DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-142-419-255 (1-1432)

QY 10 SerleuglyaspSergluThrluSergln 19

Db 1261 AGCCTGGGTGACAGCAGACCCCTGTCTCAA 1290

## RESULT 45

US-10-123-262-255

; Sequence 255, Application US/10123262

; Publication No. US20030049816A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; TITLE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3330R1C38

; CURRENT APPLICATION NUMBER: US/10/123,262

; CURRENT FILING DATE: 2002-04-15

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 255

; LENGTH: 1432

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-123-262-255

## Alignment Scores:

Pred. No.: 0.202 Length: 1432

Score: 10.00 Matches: 10

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 10.31% Indels: 0

DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-123-262-255 (1-1432)

QY 10 SerleuglyaspSergluThrluSergln 19

Db 1261 AGCCTGGGTGACAGCAGACCCCTGTCTCAA 1290

Search completed: May 11, 2003, 18:45:54  
Job time : 133.027 secs



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 11, 2003, 16:29:05 ; Search time 1473.03 Seconds  
(without alignments)  
1066.486 Million cell updates/sec

Title: US-09-854-133-586

Perfect score: 97  
Sequence: 1 EYEVSRDHASLGSETLSQT.....LTGGCLPWATRSHLGRRKCS 97

Scoring table: OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 1

Total number of hits satisfying chosen parameters: 32308013

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: listing first 60 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO\_spool/US09854133/runat\_05052003\_174133\_718/app-query.fasta\_1.462  
-DB=EST -QFMT=fastap -SUFFIX=oligo.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=60  
-DOCALIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=45 -MODE=LOCAL -OUTFMT=ptlo  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09854133@cgn2.1\_1\_2003\_runat\_05052003\_174133\_718 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlmu:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_estl1:\*  
10: gb\_estl2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	37	38.1	936 12	BG284503
2	37	38.1	1072 12	BG388107
3	33	34.0	910 12	BG326527
4	21	21.6	557 17	AO002318
5	19	19.6	123 9	AI313891
6	15	15.5	350 12	BF095483
7	13	13.4	1184 12	BF526812
8	12	12.4	357 17	AZ020264
9	12	12.4	589 17	AZ799582
10	12	12.4	629 17	AO351658
11	12	12.4	629 17	AO352329
12	12	12.4	664 17	AZ258055
13	12	12.4	871 17	AZ747681
14	12	12.4	882 12	BE880444
15	12	12.4	1032 13	BM557601
16	11	11.3	210 12	BF809298
17	11	11.3	227 12	BF743959
18	11	11.3	227 12	BF381839
19	11	11.3	259 17	AZ278638
20	11	11.3	374 14	BQ558860
21	11	11.3	382 9	AI628395
22	11	11.3	442 13	BM196350
23	11	11.3	445 17	BH117011
24	11	11.3	454 17	AZ356753
25	11	11.3	464 17	AZ779079
26	11	11.3	495 17	BH858364
27	11	11.3	500 17	BH858364
28	11	11.3	506 17	AZ750540
29	11	11.3	530 14	BQ086880
30	11	11.3	583 17	AO547131
31	11	11.3	587 14	BQ086621
32	11	11.3	643 17	AG075749
33	11	11.3	652 17	AZ791667
34	11	11.3	677 17	AG126279
35	11	11.3	689 17	AO947014
36	11	11.3	696 13	BI912987
37	11	11.3	917 12	BE733566
38	11	11.3	1024 12	BE735002
39	11	11.3	1036 13	BI912939
40	11	11.3	2909 11	BC010121
41	10	10.3	101 17	AZ742823
42	10	10.3	113 13	BI060988
43	10	10.3	127 12	BF853221
44	10	10.3	142 17	AO239603
45	10	10.3	155 10	BI122746
46	10	10.3	157 10	BB325488
47	10	10.3	158 10	BB323023
48	10	10.3	159 10	BB340999
49	10	10.3	162 10	BB182002
50	10	10.3	162 10	BB328913
51	10	10.3	162 10	BB329139
52	10	10.3	163 14	BM940420
53	10	10.3	164 10	BB322785
54	10	10.3	164 10	BB323479
55	10	10.3	165 10	BB323232
56	10	10.3	179 10	BB324227
57	10	10.3	181 10	BB323547
58	10	10.3	181 10	BB323799
59	10	10.3	182 9	AU069480
60	10	10.3		AU069480

#### ALIGNMENTS

RESULT 1  
BG284503  
LOCUS  
DEFINITION BG284503 936 bp mRNA linear EST 21-FEB-2001  
602408645F1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4537810 5',  
mRNA sequence.  
ACCESSION BG284503  
VERSION BG284503.1 GI:13035516  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 936)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: DCTD/DTF  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10463 row: 1 column: 11  
High quality sequence stop: 795.  
Location/Qualifiers  
1. 936

FEATURES  
source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4537810"  
/clone\_lib="NIH\_MGC\_91"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: prostate; Vector: PCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.4 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."  
BASE COUNT 248 a 202 c 251 g 235 t  
ORIGIN

## Alignment Scores:

Pred. No.: 2.35e-26 Length: 936  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.14% Indels: 0  
DB: 12 Gaps: 0

US-09-854-133-586 (1-97) x BG284503 (1-936)

QY 61 GluSerLeuLeuCysProProSerProLysGluValThrCysArgGluMetLeuThrGly 80  
|||||  
Db 278 GAAAGCCTGTGTGTCACCATCTCCAAAGAGGAGTTACCTGCAGGAAATGTTAACGGGA 337  
QY 81 GlyCysLeuProTrpAlaThrArgSerHisLeuGlyArgArgLysCysSer 97  
|||||  
Db 338 GGCTGCCTTCCTCGGCAACAGAGCACCCTGGCGAGAGAAAGTGCAGC 388

RESULT 2  
BG388107  
LOCUS  
DEFINITION BG388107 1072 bp mRNA linear EST 12-MAR-2001  
602413070F1 NIH\_MGC\_92 Homo sapiens cDNA clone IMAGE:4521736 5',  
mRNA sequence.  
ACCESSION BG388107  
VERSION BG388107.1 GI:13281553  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo..

REFERENCE 1 (bases 1 to 1072)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10421 row: k column: 17  
High quality sequence stop: 625.  
Location/Qualifiers  
1. 1072

FEATURES  
source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4521736"  
/clone\_lib="NIH\_MGC\_92"  
/tissue\_type="embryonal carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: testis; Vector: PCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 2.5 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."  
BASE COUNT 331 a 229 c 270 g 242 t  
ORIGIN

## Alignment Scores:

Pred. No.: 2.61e-26 Length: 1072  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.14% Indels: 0  
DB: 12 Gaps: 0

US-09-854-133-586 (1-97) x BG388107 (1-1072)

QY 61 GluSerLeuLeuCysProProSerProLysGluValThrCysArgGluMetLeuThrGly 80  
|||||  
Db 296 GAAAGCCTGTGTGTCACCATCTCCAAAGAGGAGTTACCTGCAGGAAATGTTAACGGGA 355  
QY 81 GlyCysLeuProTrpAlaThrArgSerHisLeuGlyArgArgLysCysSer 97  
|||||  
Db 356 GGCTGCCTTCCTCGGCAACAGAGCACCCTGGCGAGAGAAAGTGCAGC 406

RESULT 3  
BG326527  
LOCUS  
DEFINITION BG326527 910 bp mRNA linear EST 27-FEB-2001  
602425373F1 NIH\_MGC\_14 Homo sapiens cDNA clone IMAGE:4562994 5',  
mRNA sequence.  
ACCESSION BG326527  
VERSION BG326527.1 GI:13132964  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 910)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: DCTD/DTF  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:



http://image.llnl.gov  
Plate: L1CM1276 row: b column: 19  
High quality sequence stop: 706.  
Location/Qualifiers  
1. 910

FEATURES  
Source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4562994"  
/clone\_lib="NH\_MGC\_14"  
/tissue\_type="renal cell adenocarcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: kidney; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."  
BASE COUNT 237 a 178 c 249 g 246 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2.01e-22 Length: 910  
Score: 33.00 Matches: 33  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 34.02% Indels: 0  
DB: 12 Gaps: 0

US-09-854-133-586 (1-97) x BG326527 (1-910)

QY 65 CysProPserProlysgluValThrCysargGluMetLeuThrGlyGlyCysLeuPro 84  
LOCUS ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
Db 179 TGTCCACCATCTCCAAGAGAGGTTACCTGCAGGGAATGTTAACGGAGGCTGCTTCCC 238

QY 85 TrpAlaThrArgSerHisLeuGlyArgArgGlyCysSer 97  
Db 239 TGGGCAACAAGAGAGCCACTGGGCAAGAGAAAGTGCAC 277

RESULT 4  
AQ002318/c 557 bp DNA linear GSS 26-JUN-1998  
LOCUS CIT-HSP-2283E8.TF CIT-HSP Homo sapiens genomic clone 2283E8, DNA  
DEFINITION sequence.  
ACCESSION AQ002318  
VERSION AQ002318.1 GI:3029522  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 557)  
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,  
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,  
Simon,M. and Venter,J.C.  
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map  
BUILDING (1998)  
JOURNAL Unpublished (1998)  
COMMENT Other\_GSSs: CIT-HSP-2283E8.TR  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdadams@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
http://www.tigr.org/tdb/hungen/bac\_end\_search/bac\_end\_search.html  
Seq primer: M13-21;  
Class: BAC ends.

FEATURES  
Location/Qualifiers

source

1. .557  
/organism="Homo sapiens"  
/db\_xref="GDB:7147907"  
/db\_xref="taxon:9606"  
/clone="2283E8"  
/clone\_lib="CIT-HSP"  
/sex="Male"  
/cell\_type="Sperm"  
/note="Vector: pBelOBAC11; Site\_1: HindIII; Site\_2:  
HindIII"  
BASE COUNT 142 a 152 c 133 g 130 t  
ORIGIN

Alignment Scores:  
Pred. No.: 9.18e-11 Length: 557  
Score: 21.00 Matches: 21  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 21.65% Indels: 0  
DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x AQ002318 (1-557)

QY 72 ValThrCysArgGluMetLeuThrGlyGlyCysLeuProTrpAlaThrArgSerHisLeu 91  
Db 344 GTTACCTGCAGGGAATGTTAACGGAGGCTGCTTCCCTGGGCAACAAGAGCCACTG 285

QY 92 Gly 92  
Db 284 GGC 282

RESULT 5  
A1313891/c 123 bp mRNA linear EST 17-DEC-1998  
LOCUS hpi-7 PMA-induced HL60 cell subtraction library Homo sapiens cDNA,  
DEFINITION mRNA sequence.  
ACCESSION A1313891  
VERSION A1313891.1 GI:4029010  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 123)  
AUTHORS Hiken,J.F. and Huang,H.V.  
TITLE Genes Up-Regulated During PMA-Induced HL60 Cell Macrophage-Like  
Differentiation  
JOURNAL Unpublished (1998)  
COMMENT Contact: Hiken JF  
Department of Molecular Microbiology  
Washington University School of Medicine  
660 South Euclid Ave, Saint Louis, MO 63110-1093, USA  
Tel: 314 362 2756  
Fax: 314 362 1232  
Email: hiken@borcim.wustl.edu  
Orientation of insert unknown.  
FEATURES  
Source  
1. 123  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="PMA-induced HL60 cell subtraction library"  
/cell\_type="macrophage-like"  
/cell\_line="PMA-induced HL60 human leukemic cell line"  
/lab\_host="MC1061"  
/note="Vector: pSPORT 1; Site\_1: EcoRI; Site\_2: EcoRI; RNA  
from uninduced HL60 cells was subtracted against RNA from  
HL60 cells that were induced 8h with 16 nM PMA using the  
method of Wang and Brown (Proc Natl Acad Sci USA 88: 11505  
)."

BASE COUNT 23 a 37 c 34 g 29 t  
ORIGIN  
Alignment Scores:

Pred. No.: 2.65e-09  
Score: 19.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 19.59%  
DB: 9  
Matches: 123  
Conservative: 19  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x A1313891 (1-123)

QY 78 LeuThrglyCysLeuProTrrpAlarhArgSerHisLeuGlyArgArgLysCys 96  
|||||  
Db 59 TTAACGGAGAGCTGCTCTCCCTGGCAACAGAGCCACCTGGGAGAGAGAAAGTGC 3

RESULT 6  
BF095483

LOCUS 350 bp mRNA linear EST 19-OCT-2000  
DEFINITION IL2-UT0074-040900-152-C01 UT0074 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF095483  
VERSION BF095483.1 GI:10901193  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagal, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare  
, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

## JOURNAL

COMMENT 20202663  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=6t2-IL2-UT0074-040  
900-152-C01&t3=2000-09-04&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 24  
High quality sequence stop: 350.  
Location/Qualifiers  
1. 350

## FEATURES

## Source

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="UT0074"  
/dev\_stage="Adult"  
/note="Organ: uterus\_tumor; Vector: puc18; Site\_1: SmaI;  
Site\_2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."  
BASE COUNT 95 a 80 c 87 t  
ORIGIN

## Alignment Scores:

Pred. No.: 5.23e-05  
Score: 15.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 15.46%  
DB: 12

US-09-854-133-586 (1-97) x BF095483 (1-350)

QY 61 GluSerLeuLeuGlyCysProPserProLysGluValThrCysArg 75  
|||||  
Db 297 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGAGTTACTGCAGG 341

RESULT 7  
BF526812 1184 bp mRNA linear EST 11-DEC-2000  
LOCUS 602070364F1 NCI\_CGAP\_Brn64 Homo sapiens cDNA clone IMAGE:4213479  
DEFINITION BF526812  
5', mRNA sequence.  
ACCESSION BF526812  
VERSION BF526812.1 GI:11614175  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@psh1.nih.gov  
Tissue Procurement: David N. Louis, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LINL at:  
http://image.llnl.gov  
Plate: LAM9785 row: 0 column: 16  
High quality sequence stop: 565.  
Location/Qualifiers  
1. 1184

TITLE  
JOURNAL  
COMMENT

## JOURNAL

FEATURES  
Source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4213479"  
/clone\_lib="NCI\_CGAP\_Brn64"  
/tissue\_type="gliblastoma with EGFR amplification"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: brain; Vector: PCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.57 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."  
BASE COUNT 435 a 235 c 269 g 245 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.0126  
Score: 13.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 13.40%  
DB: 12  
Matches: 1184  
Conservative: 13  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x BF526812 (1-1184)

QY 8 HisAlaSerLeuGlyAspSerGluThrLeuSerGlnThr 20  
|||||  
Db 61 CACGCTAGCCTGGGTGACAGTGAACCTGTCTCAACA 99

## RESULT 8

AZ020264/c 357 bp DNA linear GSS 25-FEB-2000  
LOCUS  
DEFINITION RPCI-23-299J8.TV RPCI-23 Mus musculus genomic clone RPCI-23-299J8,  
DNA sequence.  
ACCESSION AZ020264  
VERSION AZ020264.1 GI:7095648  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus

Alignment Scores:  
Pred. No.: 0.0126  
Score: 13.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 13.40%  
DB: 12  
Matches: 1184  
Conservative: 13  
Mismatch: 0  
Indels: 0  
Gaps: 0

REFERENCE  
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 357)  
Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P., and Fraser, C.M.

TITLE  
JOURNAL  
COMMENT

Mouse BAC End Sequences from Library RPCI-23  
Unpublished (1999)  
Other\_GSSs: RPCI-23-299J8.TJ  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Reseach Genetics ([info@resgen.com](http://info@resgen.com)). BAC end page: [http://www.tigr.org/tldb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html)  
Plate: 299 row: J column: 8  
Seq primer: T7  
Class: BAC ends.

FEATURES  
source

location/Qualifiers  
1..357  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-23-299J8"  
/clone\_lib="RPCI-23"  
/sex="Female"  
/lab\_host="DH10B"  
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACe3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."  
49 a 82 c 27 g 194 t 5 others

## Alignment Scores:

Pred. No.: 0.048 Length: 357  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x AZ020264 (1-357)

OY 23 ArgLysLysGluArgLysLysLysArgGluArgLys 34  
|||||

Db 245 AGAAGAGAAAGAAAGAAAGAAAGAGAGAGAAAG 210

## RESULT 9

AZ799582 589 bp DNA linear GSS 16-FEB-2001  
LOCUS 2M0057F09F Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
DEFINITION clone UUGC2M0057F09 F, DNA sequence.

ACCESSION AZ799582  
VERSION AZ799582.1 GI:12950846

KEYWORDS  
SOURCE  
ORGANISM

house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 589)  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly

, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A., and Wright, D., Weiss, R.  
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL  
COMMENT

Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00  
Plate: 0057 row: F column: 09  
Seq primer: CGTTGTAACGACGCCAGT  
Class: plasmid ends  
High quality sequence stop: 589.

FEATURES  
source

location/Qualifiers  
1..589  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC2M0057F09"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: pMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g114732114/9b/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT  
ORIGIN

80 a 170 c 91 g 248 t

## Alignment Scores:

Pred. No.: 0.0709 Length: 589  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x AZ799582 (1-589)

OY 23 ArgLysLysGluArgLysLysLysArgGluArgLys 34  
|||||

Db 168 AGAAGAGAAAGAAAGAAAGAAAGAGAGAGAGAG 133

## RESULT 10

AQ351658 629 bp DNA linear GSS 24-JAN-1999  
LOCUS AQ351658 CITBI-E1-2529K22.TF CITBI-E1 Homo sapiens genomic clone 2529K22,  
DEFINITION DNA sequence.

ACCESSION AQ351658  
VERSION AQ351658.1 GI:4178993

KEYWORDS  
SOURCE  
ORGANISM

human.  
Homo sapiens

REFERENCE  
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 629)  
Zhao, S., Adams, M.D., Nierman, W., Malek, J., Shizuya, H., Simon, M. and  
Venter, J.C.

## TITLE

Use of BAC End Sequences from Caltech Libraries for Sequence-Ready  
Map Building

JOURNAL  
COMMENT

Unpublished (1997)

Other\_GSSS: CITBI-E1-2529K22.TR

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbe@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:  
[http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html).

Seq primer: M13-21

Class: BAC ends.

## FEATURES

## source

Location/Qualifiers

1..629

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="2529K22"

/clone\_lib="CITBI-E1"

/sex="male"

/cell\_type="sperm"

/note="Vector: pBelOBAC11; Site\_1: EcoRI; Site\_2: EcoRI;  
Caltech Human BAC Library D"

BASE COUNT  
ORIGIN

112 a 170 c 96 g 251 t

## Alignment Scores:

Pred. No.: 0.0746 Length: 629  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x AQ351658 (1-629)

QY 23 ArgLysLysGluArgLysLysLysArgGluArgLys 34

Db 502 AGAAGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 467

## RESULT 11

AQ352329/c

LOCUS

DEFINITION

AQ352329 629 bp DNA linear GSS 24-JAN-1999  
CITBI-E1-2529M24.TF CITBI-E1 Homo sapiens genomic clone 2529M24,

ACCESSION

AQ352329 GI:4179664

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 629)  
Zhao, S., Adams, M.D., Nierman, W., Malek, J., Shizuya, H., Simon, M. and  
Venter, J.C.  
Use of BAC End Sequences from Caltech Libraries for Sequence-Ready  
Map Building  
Unpublished (1997)  
Other\_GSSS: CITBI-E1-2529M24.TR  
Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208

FEATURES  
source

Email: hbe@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
[http://www.tigr.org/tdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html).  
Seq primer: M13-21  
Class: BAC ends.

Location/Qualifiers

1..629

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="2529M24"

/clone\_lib="CITBI-E1"

/sex="male"

/cell\_type="sperm"

/note="Vector: pBelOBAC11; Site\_1: EcoRI; Site\_2: EcoRI;  
Caltech Human BAC Library D"

BASE COUNT  
ORIGIN

112 a 169 c 96 g 252 t

## Alignment Scores:

Pred. No.: 0.0746 Length: 629  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x AQ352329 (1-629)

QY 23 ArgLysLysGluArgLysLysLysArgGluArgLys 34

Db 502 AGAAGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 467

## RESULT 12

AZ258055/c

LOCUS

DEFINITION

AZ258055 664 bp DNA linear GSS 26-JUL-2000  
RPCI-23-134E11.TV RPCI-23 Mus musculus genomic clone RPCI-23-134E11

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 664)  
Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret  
, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.  
and Fraser, C.M.  
Mouse BAC End Sequences from Library RPCI-23  
Unpublished (1999)  
Other\_GSSS: RPCI-23-134E11.TV  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the mouse BAC library RPCI-23. For BAC  
library availability, please contact Pieter de Jong  
(pieter@dejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)  
or from Resea ch Genetics (info@resgen.com). BAC end page:  
[http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
Plate: 134 row: E column: 11  
Seq primer: T7  
Class: BAC ends.

## FEATURES

## source

Location/Qualifiers

1..664

/organism="Mus musculus"

/strain="C57Bl/6J"

/db\_xref="taxon:10090"



```
/clone="RPCI-23-134E11"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
ECORI; Site_2: ECORI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcorI and EcorI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
ECORI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT      178 a      132 c      63 g      291 t
ORIGIN

Alignment Scores:
Pred. No.:      0.0778      Length:      664
Score:          12.00      Matches:      12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    12.37%      Indels:      0
DB:             17      Gaps:          0

US-09-854-133-586 (1-97) x AZ258055 (1-664)

OY      23  ArgLysLysGluArgLysLysLysArgGluArgLys 34
         |||||
Db      315 AGAAAGAAAGAAAGAAAGAAAGAGAGAGAGAGAAA 280

RESULT 13
AZ747681      871 bp      DNA      linear      GSS 25-JAN-2001
LOCUS      RPCI-24-65L15.TV RPCI-24 Mus musculus genomic clone RPCI-24-65L15,
DEFINITION      DNA sequence.
ACCESSION      AZ747681      GI:12531760
VERSION      AZ747681.1
KEYWORDS      GSS.
SOURCE      house mouse.
ORGANISM      Mus musculus.
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 871)
Zhao,S., Niemman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
Russell,D., de Jong,P. and Fraser,C.M.
Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)
Other_GSSs: RPCI-24-65L15.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@email.cho.org). Clones may be purchased from BACPAC
Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end
page: http://www.tigr.org/tdb/bac\_ends/mouse/bac\_end\_intro.html
Plate: 65 row: L column: 15
Seq primer: T7
Class: BAC ends.
FEATURES
Source      Location/Qualifiers
1..871
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-65L15"
/clone_lib="RPCI-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pTARBAC1; Site_1: BamHI; Site_2: BamHI;
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
```

```
library was cloned in the pTARBAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."
BASE COUNT      198 a      174 c      77 g      422 t
ORIGIN

Alignment Scores:
Pred. No.:      0.0961      Length:      871
Score:          12.00      Matches:      12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    12.37%      Indels:      0
DB:             17      Gaps:          0

US-09-854-133-586 (1-97) x AZ747681 (1-871)

OY      23  ArgLysLysGluArgLysLysLysArgGluArgLys 34
         |||||
Db      102 AGAAAGAAAGAAAGAAAGAAAGAGAGAGAGAGAAA 67

RESULT 14
BE880444      882 bp      mRNA      linear      EST 20-OCT-2000
LOCUS      601491430F1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3893886 5',
DEFINITION      mRNA sequence.
ACCESSION      BE880444      GI:10329220
VERSION      BE880444.1
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 882)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LAM9682 row: k column: 07
High quality sequence stop: 662.
FEATURES
Source      Location/Qualifiers
1..882
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3893886"
/clone_lib="NIH_MGC_69"
/tissue_type="large cell carcinoma, undifferentiated"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
BASE COUNT      263 a      165 c      182 g      272 t
ORIGIN

Alignment Scores:
Pred. No.:      0.0971      Length:      882
Score:          12.00      Matches:      12
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    12.37%      Indels:      0
DB:             12      Gaps:          0

US-09-854-133-586 (1-97) x BE880444 (1-882)

OY      8  HisAlaSerLeuGlyAspSerGluThrLeuSerGln 19
```

—

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 227)	Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.	Shotgun sequencing of the human transcriptome with ORF expressed

JOURNAL MEDLINE  
COMMENT

Sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM2&tl2=CM2-BT0826-181000-428-f07&tl3=2000-10-18&tl4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 32  
High quality sequence stop: 227.  
Location/Qualifiers  
1. .227  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="BT0826"  
/dev\_stage="Adult"  
/note="Organ: breast; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 36 a 64 c 37 g 90 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.326 Length: 227  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 12 Gaps: 0

US-09-854-133-586 (1-97) x BF743959 (1-227)

QY 23 ArgLysLysGluArgLysLysArgGluArg 33  
|||||  
Db 175 CGAAGAAGAGAGAGAAAGAGAGAGAGA 143

RESULT 18  
BF381839/c 227 bp mRNA linear EST 27-NOV-2000  
LOCUS 601816096F1 NIH\_MGC\_56 Homo sapiens cDNA clone IMAGE:4050286 5',  
DEFINITION mRNA sequence.  
ACCESSION BF381839  
VERSION BF381839.1 GI:11363140  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 227)  
NIH-MGC http://mgc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: CLONETECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILNL at:  
http://image.llnl.gov

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Plate: LICM883 row: 0 column: 23  
High quality sequence stop: 227.  
Location/Qualifiers  
1. .227  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4050286"  
/clone\_lib="NIH\_MGC\_56"  
/issue\_type="primitive neuroectoderm"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site\_1: SfiI (ggccgctcgcc); Site\_2: SfiI (ggccattatggc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATTCTAGAGCGCGAGCGCGGCGACATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 49 a 51 c 49 g 78 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.326 Length: 227  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 12 Gaps: 0

US-09-854-133-586 (1-97) x BF381839 (1-227)

QY 10 SerLeuGlyAspSerGluThrLeuSerGluThr 20  
|||||  
Db 77 AGCCTGGGTGACAGTGTGACTCTCTCAACA 45

RESULT 19  
A2278638/c 259 bp DNA linear GSS 26-JUL-2000  
LOCUS A2278638  
DEFINITION RPCI-23-110P5.TV RPCI-23 Mus musculus genomic clone RPCI-23-110P5,  
DNA sequence.  
ACCESSION A2278638  
VERSION A2278638.1 GI:9497540  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 259)  
Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akınret,  
B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P.  
and Fraser, C.M.  
Mouse BAC End Sequences from Library RPCI-23  
Unpublished (1999)  
Other\_GSSs: RPCI-23-110P5.TJ  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the mouse BAC library RPCI-23. For BAC  
library availability, please contact Pieter de Jong  
(pieter@dejong.med.bufileo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.bufileo.edu/orderingframe.htm)  
or from Resea ch Genetics (info@resgen.com). BAC end page:  
http://www.tigr.org/tldb/bac\_ends/mouse/bac\_end\_intro.html  
Plate: 110 row: P column: 5  
Seq primer: T7  
Class: BAC ends.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT





5'-GAGCTAGTTCTAGATCGGAGCGGGCCCTTTTTTTTTTTTT-3') from 14.2 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker L1-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon

BASE COUNT	70 a	102 c	81 g	192 t
ORIGIN				



US-09-854-133-586 (1-97) x AZ356753 (1-464)

QY 23 ArgLysLysGluArgLysLysArgGluArg 33  
|||||  
DB 405 AGAAGAAGAAAGAAAGAAAGAGAGAGA 437

## RESULT 26

AZ779079/c 495 bp DNA linear GSS 16-FEB-2001  
LOCUS 2M0014P19R Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
DEFINITION clone UUGC2M0014P19 R, DNA sequence.

ACCESSION AZ779079 GI:12909373  
VERSION  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus

## REFERENCE

AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly  
,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.  
and Wright,D., Weiss,R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunne@genetics.utah.edu  
Insert length: 10000 Std Error: 0.00  
Plate: 0014 row: P column: 19  
Seq primer: CACACAGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 495.  
Location/Qualifiers  
1. 495

JOURNAL  
COMMENT

/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC2M0014P19"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PMD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a  
0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adapted DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of PMD42 (gil4732114|gb|AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adapted mouse DNA was annealed to  
adapted vector DNA, and transformed into  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

FEATURES  
source

BASE COUNT 107 a 130 c 52 g 206 t  
ORIGIN  
Alignment Scores: 0.599 Length: 495  
Pred. No.:

Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x AZ779079 (1-495)

QY 23 ArgLysLysGluArgLysLysArgGluArg 33  
|||||  
DB 90 AGAAGAAGAAAGAAAGAAAGAGAGAGA 58

## RESULT 27

BH858364/c 500 bp DNA linear GSS 08-JUL-2002  
LOCUS B5\_164a2\_s6 Mouse Retroviral Tagged Cancer Gene Database Mus  
DEFINITION musculus genomic clone B5\_164a2, DNA sequence.

ACCESSION BH858364 GI:21709185  
VERSION  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 500)  
Suzuki,T., Shen,H., Akagi,K., Morse,H.C., Malley,J.D., Nalman,D.Q.,  
Jenkins,N.A. and Copeland,N.G.  
Retroviral tagging provides a potent cancer gene discovery tool in  
the post-genome-sequence era  
Nat. Genet., (2002) In press  
Contact: Copeland NG  
Mouse Cancer Genetics Program  
National Cancer Institute  
Bldg. 539, Rm. 229, Frederick, MD 21702-1201, USA  
Tel: 301 846 1260  
Fax: 301 846 6666  
Email: copeland@ncifcrf.gov  
Class: PCR with specific primers.  
Location/Qualifiers  
1. 500

JOURNAL  
COMMENT

/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="B5\_164a2"  
/clone\_lib="Mouse Retroviral Tagged Cancer Gene Database"  
/sex="female"  
/tissue\_type="Leukemia"  
/note="Inverse PCR method  
(http://genome2.ncifcrf.gov/RTGcd)"  
Location/Qualifiers  
1. 500

FEATURES  
source

BASE COUNT 101 a 111 c 106 g 182 t  
ORIGIN  
Alignment Scores: 0.604 Length: 500  
Pred. No.: 11.00 Matches: 11  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 11.34% Indels: 0  
Query Match: 17 Gaps: 0  
DB: 17

US-09-854-133-586 (1-97) x BH858364 (1-500)

QY 24 LysLysGluArgLysLysArgGluArgLys 34  
|||||  
DB 233 AAAAAAGAAAGAAAGAAAGAGAGAGAGAAA 201

## RESULT 28

AZ750540 506 bp DNA linear GSS 25-JAN-2001  
LOCUS RPCI-24-112119.TJ RPCI-24 Mus musculus genomic clone RPCI-24-112119  
DEFINITION , DNA sequence.

ACCESSION AZ750540  
VERSION AZ750540.1 GI:12535699  
KEYWORDS GSS.

Alignment Scores:

0.599

Length:

495

SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS 1 (bases 1 to 506)  
Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akinret, B., Levins, M.,  
Tsegaye, G., Geer, K., Krol, M., Shwartsbeyn, A., Gebregeorgis, E.,  
Russell, D., de Jong, P., and Fraser, C.M.  
Mouse BAC End Sequences from Library RPCI-24  
Unpublished (1999)  
Other\_GSSs: RPCI-24-112119.TV  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the mouse BAC library RPCI-24. For BAC  
library availability, please contact Pieter de Jong  
(pdejong@mail.cho.org). Clones may be purchased from BACPAC  
Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end  
page: [http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
Plate: 112 row: I column: 19  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
source Location/Qualifiers  
1..506  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-24-112119"  
/clone\_lib="RPCI-24"  
/sex="Male"  
/cell\_type="Spleen/Brain"  
/note="Vector: pTARBAC1; Site\_1: BamHI; Site\_2: BamHI;  
RPCI-24 Mouse BAC library produced by Pieter de Jong. The  
library was cloned in the pTARBAC1 cloning vector at the  
BamHI sites using MboI partially digested male C57BL/6J  
DNA."

BASE COUNT 232 a 69 c 131 g 74 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.609 Length: 506  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x AZ750540 (1-506)

OY 23 ArgLysLysGluArgLysLysLysArgGluArg 33  
|||||  
Db 130 AGAAGAAGAAAGAAAGAAAGAAAGAGAGAGAGA 162  
|||||

RESULT 29  
BQ086880 530 bp mRNA linear EST 05-APR-2002  
LOCUS ih91d05.y1 Melton Mouse E16 5 Pancreas library 2 M16B2 Mus musculus  
DEFINITION CDNA clone IMAGE:5939385 5', mRNA sequence.  
ACCESSION BQ086880  
VERSION BQ086880.1 GI:20046084  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 530)  
AUTHORS Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaestner, K.,  
Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S.,  
Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,

TITLE  
JOURNAL  
COMMENT  
Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas  
, M., Gibbons, M., McCann, R., Cole, R., Tsagarelshvili, R., Williams, T.,  
, Jackson, Y., and Bowers, Y.  
Endocrine Pancreas Consortium  
Unpublished (2000)  
Other\_ESTs: ih91d05.x1  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@biohp.harvard.edu  
Library was constructed by Dr. Douglas Melton DNA sequencing by:  
Washington University Genome Sequencing Center This clone is  
available royalty-free through LLNL; please contact the IMAGE  
consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information  
Seq primer: -40RP from Gibco  
High quality sequence stop: 432.

FEATURES  
source Location/Qualifiers  
1..530  
/organism="Mus musculus"  
/strain="ICR"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5939385"  
/clone\_lib="Melton Mouse E16 5 Pancreas Library 2 M16B2"  
/sex="Both"  
/tissue\_type="Total pancreas"  
/dev\_stage="Embryonic day 16.5"  
/lab\_host="TOP10"  
/note="Organ: Pancreas; Vector: pBluescript II SK; Site\_1:  
NotI; Site\_2: SalI; Library constructed using Superscript  
Plasmid Library kit (Life Technologies). CDNA made by  
oligo-dT priming. Size-selected by column fractionation;  
average insert size 1.06kb. Primary library,  
unamplified."

BASE COUNT 244 a 74 c 116 g 96 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.632 Length: 530  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 14 Gaps: 0

US-09-854-133-586 (1-97) x BQ086880 (1-530)

OY 23 ArgLysLysGluArgLysLysLysArgGluArg 33  
|||||  
Db 267 AGAAGAAGAAAGAAAGAAAGAAAGAGAGAGAGA 299  
|||||

RESULT 30  
AQ547131 583 bp DNA linear GSS 28-MAY-1999  
LOCUS AQ547131  
DEFINITION RPCI-11-431G8.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-431G8,  
DNA sequence.  
ACCESSION AQ547131  
VERSION AQ547131.1 GI:4906374  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 583)  
AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter  
, J.C.  
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
Map Building  
Unpublished (1997)  
Other\_GSSs: RPCI-11-431G8.TV



Contact: Shaying Zhao, William Nierman, Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850  
 Tel: 301 838 0200  
 Fax: 301 838 0208

Email: hbeetigr.org  
 Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@edj.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genet cs (<http://info@resgen.com>). BAC end search page: [http://www.tigr.org/tldb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html).  
 Seq primer: SP6  
 Class: BAC ends.

# FEATURES

source location/Qualifiers  
 1. 583  
 /organism="Homo sapiens"  
 /db\_xref="GDB:7665271"  
 /db\_xref="taxon:9606"  
 /clone="RPCI-11-431G8"  
 /clone\_lib="RPCI-11"  
 /sex="Male"  
 /cell\_type="Lymphocytes"  
 /note="Vector: PBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI; RPC11 Human Male BAC library"

BASE COUNT 81 a 130 c 50 g 322 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.68 Length: 583  
 Score: 11.00 Matches: 11  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 11.34% Indels: 0  
 DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x AG0547131 (1-583)

QY 24 LysLysGluArgLysLysLysArgGluArgLys 34  
 |||||  
 Db 274 AAGAAAGAAAGAAAGAAAGAGAGAAAGAAAG 242

## RESULT 31

BQ086621 587 bp mRNA linear EST 05-APR-2002  
 LOCUS 1h91d05.x1 Melton Mouse E16 5 Pancreas Library 2 M16B2 Mus musculus  
 DEFINITION cdna clone IMAGE:5939385 3', mRNA sequence.

ACCESSION BQ086621  
 VERSION BQ086621.1 GI:20045825

KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 587)  
 Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hiller,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarisvilli,R., Williams,T., Jackson,Y. and Bowers,Y.  
 Endocrine Pancreas Consortium  
 Unpublished (2000)  
 Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138

TITLE JOURNAL  
 COMMENT

Tel: 617-495-1812  
 Fax: 617-495-8557  
 Email: dmelton@biohp.harvard.edu  
 Library was constructed by Dr. Douglas Melton DNA sequencing by:

Washington University Genome Sequencing Center This clone is available royalty-free through LNL; please contact the IMAGE consortium ([info@image.lnl.gov](http://info@image.lnl.gov)) for further information  
 Seq primer: -40UP from G1bco  
 High quality sequence stop: 466.

# FEATURES

source location/Qualifiers  
 1. 587  
 /organism="Mus musculus"  
 /strain="ICR"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:5939385"  
 /clone\_lib="Melton Mouse E16 5 Pancreas Library 2 M16B2"  
 /sex="Both"  
 /tissue\_type="Total pancreas"  
 /dev\_stage="Embryonic day 16.5"  
 /lab\_host="TOP10"  
 /note="Organ: Pancreas; Vector: pBluescript II SK; Site\_1: NotI; Site\_2: SalI; Library constructed using Superscript Plasmid Library kit (Life Technologies). cDNA made by oligo-dT priming. Size-selected by column fractionation; average insert size 1.06kb. Primary library, unamplified."

BASE COUNT 170 a 106 c 88 g 223 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0.684 Length: 587  
 Score: 11.00 Matches: 11  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 11.34% Indels: 0  
 DB: 14 Gaps: 0

US-09-854-133-586 (1-97) x BQ086621 (1-587)

QY 23 ArgLysLysGluArgLysLysLysArgGluArg 33  
 |||||  
 Db 564 AGAAGAAAGAAAGAAAGAAAGAGAGAGAGA 532

## RESULT 32

AG075749 643 bp DNA linear GSS 03-NOV-2001  
 LOCUS Pan troglodytes DNA, clone: PTB-068L23.F, genomic survey sequence.  
 DEFINITION AG075749  
 ACCESSION AG075749  
 VERSION AG075749.1 GI:16627551

KEYWORDS GSS.  
 SOURCE Pan troglodytes male lymphoblast DNA, clone\_lib:PTB Chimpanzee Male BAC library clone:PTB-068L23.F.

ORGANISM Pan troglodytes

REFERENCE 1  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE BAC end sequences of Library PTB  
 JOURNAL Unpublished

REFERENCE 2 (bases 1 to 643)  
 Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

TITLE Direct Submission  
 JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou,Tsukumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:chimpbes@sc.riken.go.jp, URL:<http://hgp.gsc.riken.go.jp/>, Tel:81-45-503-9111, Fax:81-45-503-9170)  
 Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.  
 PRIMERS  
 Sequencing: -21M13  
 LIBRARY  
 Vector : pRS145  
 R.Site 1 : SacI

FEATURES R.Site 2 : SacI.  
Location/Qualifiers  
1. .643  
/organism="Pan troglodytes"  
/db\_xref="taxon:9598"  
/clone="PTB-068L23.F"  
/sex="male"  
/cell\_type="lymphoblast"  
/clone\_lib="PTB Chimpanzee Male BAC Library"  
ORIGIN 165 a 144 c 128 g 205 t 1 others

Alignment Scores:  
Pred. No.: 0.734 Length: 643  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x AG075749 (1-643)

QY 24 LysLysGLuArgLysLysLysArgLys 34  
|||||  
Db 566 AAGAAAGAAAGAAAAAGAGAGAGAGAGAG 534

RESULT 33  
AZ791667 652 bp DNA linear GSS 16-FEB-2001  
LOCUS 2M0041K09 Mouse 10kb plasmid UUGC1M library Mus musculus genomic  
ACCESSION AZ791667  
VERSION AZ791667.1 GI:12934796  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 652)  
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,  
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.  
and Wright,D., Weils,R.  
TITLE Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
JOURNAL Unpublished (2000)  
COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0041 row: K column: 09  
Seq primer: CACACAGGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 652.  
Location/Qualifiers  
1. .652  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUGC2M0041K09"  
/clone\_lib="Mouse 10kb plasmid UUGC1M library"  
/sex="male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PWD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a.

FEATURES  
source

0.005 inch orifice at constant velocity. The sheared DNA  
was blunt end-repaired with T4 DNA polymerase and T4  
polynucleotide kinase. Adaptor oligonucleotides were  
ligated to the blunt ends in high molar excess. The  
adaptor DNA was purified and size-selected for a 9.5 to  
10.5 kb range using preparative agarose gel  
electrophoresis. Vector DNA was prepared from a derivative  
of PWD42 (g14732114/gb/AF129072.1), a copy-number  
inducible derivative of plasmid R1. The vector was ligated  
with adaptors complementary to the insert adaptors and  
purified. The sheared, adaptor mouse DNA was annealed to  
chemically-competent E. coli XL10-Gold (Stratagene) cells  
and selected for ampicillin resistance."

BASE COUNT 290 a 94 c 121 g 146 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 0.742 Length: 652  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x AZ791667 (1-652)

QY 23 ArgLysLysGLuArgLysLysLysArgLys 33  
|||||  
Db 368 AGAAAGAAAGAAAGAAAGAGAGAGAGAGAG 400

RESULT 34  
AG126279 677 bp DNA linear GSS 04-NOV-2001  
LOCUS AG126279/c  
DEFINITION Pan troglodytes DNA, clone: PTB-136M16.F, genomic survey sequence.  
ACCESSION AG126279  
VERSION AG126279.1 GI:16655444  
KEYWORDS GSS.  
SOURCE Pan troglodytes male lymphoblast DNA, clone\_lib:PTB Chimpanzee Male  
BAC Library clone:PTB-136M16.F.  
ORGANISM Pan troglodytes  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
1  
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,  
Totoki,Y., Watanabe,H. and Sakaki,Y.  
TITLE BAC end sequences of Library PTB  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 677)  
AUTHORS Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,  
Totoki,Y., Watanabe,H. and Sakaki,Y.  
TITLE Direct Submission  
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
1-7-22 Suenhiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail:chimbesc@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,  
Tel:81-45-503-9111, Fax:81-45-503-9170)  
Clones are derived from the chimpanzee BAC library PTB This BAC end  
was generated during the Rad process and may have higher chance of  
clone tracking errors.  
PRIMERS  
Sequencing: -21M13  
LIBRARY  
Vector : PKS145  
R.Site 1 : SacI  
R.Site 2 : SacI.  
Location/Qualifiers  
1. .677  
/organism="Pan troglodytes"  
/db\_xref="taxon:9598"  
/clone="PTB-136M16.F"  
/sex="male"  
/cell\_type="lymphoblast"

FEATURES  
source

```

BASE COUNT      169 a      220 c      153 g      135 t
ORIGIN
Alignment Scores:
Pred. No.:      0.764      Length:      677
Score:          11.00      Matches:      11
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    11.34%      Indels:      0
DB:             17      Gaps:          0

US-09-854-133-586 (1-97) x AG126279 (1-677)

QY      10 SerleuGlyAspSerGluThrLeuSerGlnThr 20
         |||||
Db       416 AGCCTGGGGACAGACGACCCCTGCTCAACA 384

RESULT 35
AC947014      689 bp      DNA      linear      GSS 27-JAN-2000
LOCUS      Sheared DNA-45F24.TF Sheared DNA Trypanosoma brucei genomic clone
DEFINITION      Sheared DNA-45F24, DNA sequence.
ACCESSION      AQ947014
VERSION        AQ947014.1  GI:6770279
KEYWORDS      GSS.
SOURCE        Trypanosoma brucei.
ORGANISM      Trypanosoma brucei.
              Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae;
REFERENCE      1 (bases 1 to 689)
AUTHORS        El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C.,
              Gerrard,C., Leech,V., de Jong,P., Ullu,E., Melville,S., Donelson,J.,
              Fraser,C. and Adams,M.
              Determination of clone end sequences from Trypanosoma brucei GUTat
              10.1 sheared DNA library
              Unpublished (1999)
              Other_GSSs: Sheared DNA-45F24.TR
              Contact: Najib M. El-Sayed
              Department of Eukaryotic Genomics
              The Institute for Genomic Research
              9712 Medical Center Dr., Rockville, MD 20850, USA
              Tel: 301 838 0200
              Fax: 301 838 0208
              Email: nelsayed@tigr.org
              Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared
              DNA library constructed at TIGR. Clones will be available for
              distribution through ATCC. Sheared DNA end sequences search page:
              http://www.tigr.org/tldb/mdb/tbdb/.
              Seq primer: M13-Forward
              Class: Shotgun.
FEATURES
source
Location/Qualifiers
1..689
/organism="Trypanosoma brucei"
/strain="TREU927/4 GUTat 10.1"
/db_xref="taxon:5691"
/clone="Sheared DNA-45F24"
/clone_lib="Sheared DNA"
/note="Vector: pUC18; Site_1: SmaI; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from a cloned population of
Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically
sheared to give a tight size distribution (approx 2 kb).
The v + l method used for the library construction is
described in detail in Smith, H.O. and Venter, J.C.
(Making small insert libraries for whole genome shotgun
sequencing projects. In Genome Sequencing: A Practical
Approach, eds. M. Vaudin and B. Borell, Oxford University
Press, 1999)."
BASE COUNT      254 a      139 c      129 g      167 t
ORIGIN
Alignment Scores:

```

```

Pred. No.:      0.775      Length:      689
Score:          11.00      Matches:      11
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    11.34%      Indels:      0
DB:             17      Gaps:          0

US-09-854-133-586 (1-97) x AQ947014 (1-689)

QY      23 ArgLysLysGluArgLysLysArgGluArg 33
         |||||
Db       94 AGAAAGAAAGAAAGAAAGAAAGAGAGAGAGA 126

RESULT 36
BI912987      696 bp      mRNA      linear      EST 16-OCT-2001
LOCUS      603176482F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5240780 5',
DEFINITION      mRNA sequence.
ACCESSION      BI912987
VERSION        BI912987.1  GI:16177268
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 696)
AUTHORS        NIH-MGC http://mgi.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgaphs-remail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              plate: L1AM11607 row: c column: 21
              High quality sequence start: 3
              High quality sequence stop: 632.
FEATURES
source
Location/Qualifiers
1..696
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5240780"
/clone_lib="NIH_MGC_121"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
fetal brains, female age 20 weeks, female age 24 weeks,
and male age 26 weeks. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
0.7-3.5 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 017. Note:
this is a NIH MGC Library."
BASE COUNT      186 a      144 c      184 g      182 t
ORIGIN
Alignment Scores:
Pred. No.:      0.781      Length:      696
Score:          11.00      Matches:      11
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    11.34%      Indels:      0
DB:             13      Gaps:          0

US-09-854-133-586 (1-97) x BI912987 (1-696)

QY      10 SerleuGlyAspSerGluThrLeuSerGlnThr 20
         |||||

```

Db 80 AGCCTGGCGACAGTGAGACCCCTGTCTCAACA 112

RESULT 37  
BE733566/c  
LOCUS  
DEFINITION BE733566 917 bp mRNA linear EST 15-SEP-2000  
601565919F1 NIH\_MGC\_21 Homo sapiens cDNA clone IMAGE:3840749 5',  
mRNA sequence.  
ACCESSION BE733566  
VERSION BE733566  
KEYWORDS BE733566.1 GI:10147558  
EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 917)  
NIH-MGC http://mhc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: LCM531 row: e column: 06  
High quality sequence stop: 677.  
Location/Qualifiers  
1. 917  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3840749"  
/clone\_lib="NIH\_MGC\_21"  
/tissue\_type="choriocarcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: placenta; Vector: pOTB7; Site\_1: XhoI;  
Site\_2: EcoRI; CDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Size-selected  
for average insert size 1.8kb. Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-CDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 213 a 201 c 244 g 258 t 1 others

ORIGIN

Alignment Scores:  
Pred. No.: 0.968 Length: 917  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 12 Gaps: 0

US-09-854-133-586 (1-97) x BE733566 (1-917)

OY 23 ArgLysLysGluArgLysLysArgGluArg 33  
|||||  
Db 63 CGAAAGAAAGAGAGAAAGAAAGAGAGAGAGA 31

RESULT 38  
BE735002/c  
LOCUS  
DEFINITION BE735002 1024 bp mRNA linear EST 15-SEP-2000  
601565360F1 NIH\_MGC\_21 Homo sapiens cDNA clone IMAGE:3840062 5',  
mRNA sequence.  
ACCESSION BE735002  
VERSION BE735002  
KEYWORDS BE735002.1 GI:10148981  
EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1024)  
AUTHORS NIH-MGC http://mhc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: LCM529 row: h column: 15  
High quality sequence start: 23  
High quality sequence stop: 781.  
Location/Qualifiers  
1. 1024  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3840062"  
/clone\_lib="NIH\_MGC\_21"  
/tissue\_type="choriocarcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: placenta; Vector: pOTB7; Site\_1: XhoI;  
Site\_2: EcoRI; CDNA made by oligo-dT priming.  
Directionally cloned into EcoRI/XhoI sites using the  
following 5' adaptor: GGCACGAG(G). Size-selected  
for average insert size 1.8kb. Library constructed by  
Ling Hong in the laboratory of Gerald M. Rubin (University  
of California, Berkeley) using ZAP-CDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 235 a 235 c 266 g 287 t 1 others

ORIGIN

Alignment Scores:  
Pred. No.: 1.05 Length: 1024  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 12 Gaps: 0

US-09-854-133-586 (1-97) x BE735002 (1-1024)

OY 23 ArgLysLysGluArgLysLysArgGluArg 33  
|||||  
Db 117 CGAAAGAAAGAGAGAAAGAAAGAGAGAGAGA 85

RESULT 39  
BI912939  
LOCUS  
DEFINITION BI912939 1036 bp mRNA linear EST 16-OCT-2001  
603176781F1 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5240756 5',  
mRNA sequence.  
ACCESSION BI912939  
VERSION BI912939  
KEYWORDS BI912939.1 GI:16177219  
EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1036)  
NIH-MGC http://mhc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov



Plate: LLAM11607 row: b column: 21  
High quality sequence start: 6  
High quality sequence stop: 158.  
Location/Qualifiers  
1. .1036

FEATURES  
source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5240756"  
/clone\_lib="NIH\_MGC\_121"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH\_MGC Library."  
BASE COUNT 244 a 234 c 259 g 294 t 5 others  
ORIGIN

Alignment Scores:  
Pred. No.: 1.06 Length: 1036  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 13 Gaps: 0

US-09-854-133-586 (1-97) x B1912939 (1-1036)

QY 10 SerLeuGLYAspSerGIuThrLeuSerGIuThr 20  
|||||  
Db 77 AGCCTGGGCGACAGTGAACCTGCTCAACA 109

RESULT 40  
BC010121/c 2909 bp mRNA linear HTC 12-JUL-2001  
LOCUS Homo sapiens, Similar to hypothetical protein MGC5149, clone  
DEFINITION IMAGE:3840062, mRNA.  
ACCESSION BC010121 GI:14714431  
VERSION BC010121.1  
KEYWORDS HTC.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
AUTHORS 1 (bases 1 to 2909)  
TITLE Strausberg, R.  
JOURNAL Direct Submission  
COMMENT Submitted (02-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaps-remail.nih.gov](mailto:cgaps-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@hgrl.nih.gov](mailto:nisc_mgc@hgrl.nih.gov)  
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Mastella, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantrilop, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,

REMARK  
COMMENT

Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAL Plate: 27 Row: 1 Column: 10  
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein  
This clone has the following problem: frame shifted.

FEATURES  
source  
Location/Qualifiers  
1. .2909  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3840062"  
/tissue\_type="Placenta, choriocarcinoma"  
/clone\_lib="NIH\_MGC\_21"  
/lab\_host="DH10B-R"  
/note="Vector: pOTB7"  
BASE COUNT 803 a 630 c 662 g 814 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2.38 Length: 2909  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 11 Gaps: 0

US-09-854-133-586 (1-97) x BC010121 (1-2909)

QY 23 ArgLYSLysGLuArgLYSLysArgGLuArg 33  
|||||  
Db 93 CGAAGAAGAGAGAGAAGAAAGAGAGAGAGA 61

RESULT 41  
AZ742823/c 101 bp DNA linear GSS 25-JAN-2001  
LOCUS RPCI-24-158G7.TJ RPCI-24 Mus musculus genomic clone RPCI-24-158G7,  
DEFINITION DNA sequence.  
ACCESSION AZ742823  
VERSION AZ742823.1 GI:12522079  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS 1 (bases 1 to 101)  
Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akinret, B., Levins, M., Tsegaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregorgis, E., Russell, D., de Jong, P., and Fraser, C.M.  
Mouse BAC End Sequences from Library RPCI-24  
Unpublished (1999)  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: [szhao@tigr.org](mailto:szhao@tigr.org)  
Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong ([pdejong@mail.cho.org](mailto:pdejong@mail.cho.org)). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end page: [http://www.tigr.org/tldb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html)  
Plate: 158 row: G column: 7  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
source

Location/Qualifiers  
1. .101  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"

```
/clone="RPCI-24-158G7"
/clone_lib="RPCI-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pTARBAC1; site_1: BamH1; site_2: BamH1;
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamH1 sites using MboI partially digested male C57BL/6J
DNA."
BASE COUNT      10 a      37 c      1 g      53 t
ORIGIN

Alignment Scores:
Pred. No.:      1.68      Length:      101
Score:          10.00      Matches:      10
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      10.31%      Indels:      0
DB:              17      Gaps:      0

US-09-854-133-586 (1-97) x AZ742823 (1-101)

QY      24      LysLysGLuArgLysLysArgGLuArg      33
Db      74      AAGAAAGAAAGAAAGAAAGAGAGAGAG      45

RESULT 42
LOCUS      BI060988      113 bp      mRNA      linear      EST 15-JUN-2001
DEFINITION      IL3-UT0116-020201-467-F11_1 UT0116 Homo sapiens cDNA, mRNA
sequence.
ACCESSION      BI060988
VERSION      BI060988.1      GI:14468515
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 113)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3&tl2=IL3-UT0116-
020201-467-F11_1&tl3=2001-02-02&tl4=1)
Seq primer: puc 18 forward
High quality sequence stop: 113.
location/Qualifiers
1.113
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="UT0116"
/dev_stage="Adult"
/note="Organ: uterus_tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
```

```
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      65 a      21 c      16 g      11 t
ORIGIN

Alignment Scores:
Pred. No.:      1.83      Length:      113
Score:          10.00      Matches:      10
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      10.31%      Indels:      0
DB:              13      Gaps:      0

US-09-854-133-586 (1-97) x BI060988 (1-113)

QY      10      SerLeuGLyAspSerGLuThrLeuSerGln      19
Db      22      AGCCTGGGTGACAGTGAGACTCTGTACCAA      51

RESULT 43
LOCUS      BF853221      127 bp      mRNA      linear      EST 16-JAN-2001
DEFINITION      MR2-EN0091-191200-002-g09_1 EN0091 Homo sapiens cDNA, mRNA
sequence.
ACCESSION      BF853221
VERSION      BF853221.1      GI:12240965
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 127)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&tl2=MR2-EN0091-
191200-002-g09_1&tl3=2000-12-19&tl4=1)
Seq primer: puc 18 forward
High quality sequence stop: 67.
location/Qualifiers
1.127
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="EN0091"
/dev_stage="Adult"
/note="Organ: lung_normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      38 a      29 c      30 g      30 t
ORIGIN
```

Alignment Scores:

Pred. No.:	2.01	Length:	127
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	12	Gaps:	0

US-09-854-133-586 (1-97) x BF853221 (1-127)

QY 10 SerleuglyaspserglutThrleuSerGln 19  
 Db 24 AGCCTGGGTGATAGCGAGACTCTGTCCAA 53

RESULT 44  
 AO239603/c  
 LOCUS  
 DEFINITION CTT-HSP-2387119.TF.1 CTT-HSP Homo sapiens genomic clone 2387119,  
 DNA sequence.

ACCESSION AO239603  
 VERSION AO239603.1 GI:3671894  
 KEYWORDS GSS.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS  
 1 (bases 1 to 142)  
 Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,  
 Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and  
 Venter,J.C.

TITLE  
 JOURNAL  
 COMMENT  
 Use of a random human BAC End Sequence Database for Sequence-Ready  
 Map Building  
 Unpublished (1998)  
 Contact: Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: mdadams@tigr.org  
 Clones are available from Research Genetics (info@resgen.com). BAC  
 end search page:  
 http://www.tigr.org/tbdb/humgen/bac\_end\_search/bac\_end\_search.html.  
 Seq primer: M13-21  
 Class: BAC ends.

FEATURES  
 source  
 Location/Qualifiers  
 1..142  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="CTT-HSP"  
 /sex="Male"  
 /cell\_type="Sperm"  
 /note="Vector: pBelobAC11; site\_1: HindIII; site\_2:  
 HindIII"

BASE COUNT 29 a 30 c 32 g 51 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2.19 Length: 142  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0  
 DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x AO239603 (1-142)

QY 10 SerleuglyaspserglutThrleuSerGln 19  
 Db 106 AGCCTGGGTGATAGCGAGACTCTGTCCAA 77

RESULT 45

BB122746/c

LOCUS BB122746 155 bp mRNA linear EST 27-JUN-2000  
 DEFINITION BB122746 RIKEN full-length enriched, adult male urinary bladder Mus  
 musculus cDNA clone 9530089K22 3', mRNA sequence.

ACCESSION BB122746  
 VERSION BB122746.1 GI:8775314  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE  
 AUTHORS  
 1 (bases 1 to 155)  
 Konno,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci  
 ,P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,  
 Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,  
 Izawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,  
 Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusakabe,M.,  
 Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,  
 Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata  
 ,Y., Shigemoto,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.,  
 Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toya  
 ,T., Tsunoda,Y., Watabiki,A., Watanabe,S., Yamamura,T., Yamanaka,I.,  
 Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino  
 ,M., Muramatsu,M. and Hayashizaki,Y.

TITLE  
 JOURNAL  
 COMMENT  
 RIKEN Mouse ESTs (Konno,H., et al.)  
 Unpublished (2000)  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center(GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,  
 URL:http://genome.gsc.riken.go.jp/  
 Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoka,S., Sasaki  
 ,N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 Thermostabilization and thermoinactivation of thermolabile enzymes by  
 trehalose and its application for the synthesis of full length  
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
 Itoh,M., Katsunuma,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,  
 Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki  
 ,Y. and Hayashizaki,Y.  
 Automated filtration-based high-throughput plasmid preparation  
 system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci,P. and Hayashizaki,Y.  
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,  
 19-44 (1999)  
 Please visit our web site (http://genome.rtc.riken.go.jp) for  
 further details.

FEATURES  
 source  
 Location/Qualifiers  
 1..155  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone\_lib="9530089K22"  
 /clone\_lib="RIKEN full-length enriched, adult male urinary  
 bladder"  
 /sex="male"  
 /tissue\_type="urinary bladder"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Site\_1: SalI; Site\_2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5',  
 GAGAGAGAGAGATCCAGAGCTTTTCTTTTCTTTTCTTTT 3'], cDNA was  
 prepared by using trehalose thermo-activated reverse  
 transcriptase and subsequently enriched for full-length by  
 cap-trapper. cDNA went through one round of normalization  
 to Rot = 20.0 and subtraction to Rot = 370.0. Second

strand cDNA was prepared with the primer adapter of  
sequence 5' GAGAGAGAGATTCGAGTTAATTATCCCCCCCCCC  
3'. cDNA was cleaved with XhoI and BamHI. Vector: a  
modified pBluescript KS(+) after bulk excision from  
Lambda FLIC I."

BASE COUNT 36 a 29 c 32 g 58 t  
ORIGIN

Alignment Scores:

Pred. No.:	2.35	Length:	155
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	10	Gaps:	0

US-09-854-133-586 (1-97) x BBI22746 (1-155)

OY 24 LysLysGluArgLysLysLysArgGluArg 33

Db 134 AAAAAAAAAAGAGAGAGAGA 105

Search completed: May 11, 2003, 17:46:15  
Job time : 1486.03 secs